# Dynamics and Representation in the Primary Visual Cortex

vorgelegt von
Diplom-Informatiker

**Péter Adorján**

Vom Fachbereich 13 - Informatik
der Technische Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
- Dr. rer. nat.-

genehmigte Dissertation

Promotionsausschuß:
Vorsitzender: Prof. Dr. Günter Hommel
Berichter: Prof. Dr. Klaus Obermayer
Berichter: Prof. Dr. Andreas V. M. Herz
Berichter: Prof. Dr. Jack D. Cowan

Tag der wissenschaftlichen Aussprache: 14.12.2000

Berlin 2000

D 83

*Grau, teurer Freund, ist alle Theorie,*
*Und grün des Lebens goldner Baum.*
(Johann Wolfgang von Goethe)

# Acknowledgments

I would like to express my greatest gratitude to Prof. Klaus Obermayer, the supervisor of this research work. Prof. Obermayer has set up an excellent laboratory, the Neural Information Processing Group at Technical University Berlin with an enormous work and attention. This free and intellectually moving scientific environment was highly inspiring for my studies. The numerous discussions, seminars, invited guests, workshops, and meetings with our collaborators were essential sources of motivating new ideas.

Among our collaborators I would like express special thanks to Prof. Jenny Lund. Her crystal clear explanations provided deep insight into the anatomy of the visual cortex. Constructive discussions with Prof. Jack D. Cowan and with Prof. Jonathan Levitt were similarly highly motivating.

I am greatly indebted to all my colleagues at the Neural Information Processing Group. It was great fun to work together in a close and fruitful collaboration with Christian Piepenbrock and Lars Schwabe. It was important synergic drive to change ideas with Cornelius Weber. The exciting conversations with Thore Graepel are unforgettable. I enjoyed several discussions with Drs. Martin Stetter, Ute Bauer and Michael Scholz.

I would like to also thank to all those people who have helped in building the foundations of my scientific work. I remember with a great pleasure the life and study in the highly pluralistic atmosphere of the Trefort Ágoston High School in Budapest. I have gained essential amount of knowledge of biologically detailed mathematical description of neural systems in Péter Érdi's laboratory (Dept. of Biophysics, Hungarian Academy of Sciences), where I could work together with Drs. György Barna and Tamás Grőbler.

I am very grateful to my grandparents and parents for their invaluable support throughout my course.

I would like to dedicate this work with all my love to Erika.

This work was supported by the German Science Foundation (Ob 102/2-1.; Gratuiertenkolleg "Signalketten in Lebenden Systems" GK 120-2), by MRC G9408137, by HFSPO RG-98/94, by Viprom Biomed2 EC grant, by Wellcome Trust 050080/Z/97 and by the VW Foundation I/71945.

# Abstract

We investigate the processing and representation of static visual patterns in the early visual system of mammals (especially cats and primates). We demonstrate that neurophysiological and anatomical findings can motivate theoretical considerations about the neural processing and vice versa. We explore *"How?"* and *"Why?"* questions in a close connection to each other. Methodologically this means using biologically detailed "bottom-up" computational models and abstract "top-down" models in parallel or in combination. Specifically, we focus on the contrast- and orientation-processing in the primary visual cortex (V1) with a strong emphasis on the dynamics of the neural activity and synapses. We consider neural dynamics on three different time scales: *(i)* the fast time evolution of the cortical activity with a time constant of $16-20$ msec; *(ii)* the intermediate modulation of the recurrent cortical competition strength with a time constant in the order of $100-200$ msec (the approximate length of a fixation period); *(iii)* contrast adaptation by the slow modulation of the dynamic nature of the synaptic transmission with a time constant of $5-10$ sec.

*Firstly*, we explore **how** orientation selectivity could be generated in the primary visual cortex (V1) (chapters 2, 3). Orientation selectivity is a remarkable and well-explored feature of the simple cells in V1. However, there is still considerable debate about the neurophysiological and anatomical origin of the highly feature selective response of these cells. The major question concerns the extent to which the simple cell properties are determined by the structure of their *feed-forward* connectivity *versus* the *recurrent* projections. In contrast to previous models, in which the initial orientation bias is generated by convergent geniculate (feed-forward) input to the simple cells, and subsequently sharpened by the lateral circuits, our approach is based on anisotropic intracortical excitatory connections. We study the hypothesis that these recurrent projections provide *both* the initial orientation bias and its subsequent amplification and therefore orientation selectivity is generated purely intracortically. Our computational study shows that indeed the "intracortical hypothesis" is a plausible alternative to the other existing hypotheses. The model predicts that the dynamics of the orientation tuning could be indicative of the underlying neural mechanism. Therefore we investigate recurrent dynamics in a cortical orientation hypercolumn in a more biologically detailed statistical neural field model (chapter 3).

*Secondly*, we study **why** the recurrent cortical re-processing of the feed-forward input is important for the representation of the image projected on the retina (chapter 4). We propose that the recurrent lateral connections implement competition between orientation selective simple cells with overlapping receptive fields. Then, we introduce the concept of *"dynamic coding"*, and investigate the short term dynamics of the recurrent competition in the primary visual cortex in terms of information processing. We find that information transfer is optimal in any increasing time window after stimulus onset if the recurrent cortical amplification decreases. In the model, the initially strong cortical competition decreases, and the role of the geniculate origin feed-forward projections becomes more important. These geniculo-cortical projections carry a topographic representation of the image projected to the retina. Motivated by information theory, our results offer a compromise between the "feed-forward" and the "recurrent" hypotheses for orientation selectivity. We suggest that both are valid, however, in different phases of the cortical processing during a fixation period. In the *initial phase* of processing, the recurrent competition is strong, and the salient orientation is signaled in a winner-take-all fashion. In the *second phase*, cortical competition becomes weaker, allowing the detection of multiple orientations. A detailed computational model provides experimentally testable

predictions about the dynamics of cortical response to multiple orientations.

*Thirdly*, we study **how** and **why** contrast adaptation occurs in V1 (chapter 5). We find that the adaptation of the *transmitter release probability* accounts well for all the puzzling experimental data that is available about the neurophysiology of contrast adaptation. The good match between our simulation results and the experimental data originates from the fact that the dynamic nature of the synaptic transmission depends on the transmitter release probability. The adaptation rule for the transmitter release probability is derived from the assumed functional objective of contrast adaptation. We propose that contrast adaptation reduces the redundancy in the cortical response by matching the activation function of single cortical neurons to the second-order signal statistics. We also show that increasing the release probability in a low-contrast environment has the functional advantage that it induces a cortical neuron to detect *synchrony* in its presynaptic spike trains, rather than the presynaptic firing *rates*. This synchrony detection mode may be proper for noise filtering if the contrast level is decreased because synchronous geniculate firing events are more likely to be stimulus related.

# Contents

# Chapter 1

# Introduction

## 1.1 Why the primary visual cortex?

The primary visual cortex (abbreviated as V1, and also referred to as area 17 or striate cortex) is the first *cortical* area that processes the visual signals arriving from the retina. The feed-forward input to V1 is preprocessed by the retinal ganglion and the relay cells of the lateral geniculate nucleus (LGN). The subcortical preprocessing in the retina and in the LGN includes detection of short scale spatial and temporal changes in the illumination pattern, detection of color, and adaptation to simple statistical properties of the visual world, like the mean luminance level.[1]

Retinal ganglion and geniculate cells have localized *receptive fields* with a shape similar to a Mexican hat. Two types of retinal ganglion and geniculate cells can be distinguished based on their receptive field profiles. The ON-center cells can be excited by illuminating the center part, and can be suppressed by illuminating the annulus surrounding their receptive field center. The OFF-center cells exhibit the opposite behavior. As a consequence of this center-opponent organization of their receptive field profiles, these cells are mainly sensitive to local changes in the illumination pattern. Diffuse illumination evokes almost no response. The activity pattern of the retinal ganglion cells is mapped topographically to the primary visual cortex via the relay cells of the LGN.

The subcortical preprocessing of the visual signals hardly involves feature extraction. Information theoretical studies propose that retinal ganglion cells eliminate simple and irrelevant—first and second order—spatial and temporal redundancies from the sensory signal to obtain a compact but sufficiently information rich representation of the visual input (Shapley and Enroth-Cugell, 1984; Atick and Redlich, 1990; Atick and Redlich, 1992). By lower order redundancies we mean the average intensity level (e.g., lightness), variance (e.g., contrast) or correlations between two points in the visual field (reflected in an unequal power spectrum).[2] In other words, the task of these subcortical preprocessing regions is to transmit maximal (or a minimally required) amount of information via their limited channel bandwidth. Transmitting maximal information on an information bottleneck enforces an efficient neural code that does not contain irrelevant or "boring" messages. This code forms the basis for the extraction of the relevant and "interesting" content in the sensory input at

[1] For a detailed description of the anatomical structure and the physiology of the subcortical visual pathways see Kandel et al. (1991).

[2] For further discussion of redundancy reduction see section 1.2 and, e.g., Atick and Redlich (1992).

later cortical processing stages.

As a result of the subcortical preprocessing, the primary visual cortex receives a faithful and ecological representation of the image *pattern*. The image pattern helps to recognize objects, shapes and therefore it carries relevant information. The pattern is formed by the higher-order correlations, i.e. constellations of different modalities between several points in the visual field or in time. The word "ecological" here refers to the very important concept that the sensory processing systems are not generally optimal, but they are adapted to our surrounding world and they also fit to our internal needs. It is also suggested that the cortical representation is faithful: in contrast to the preceding subcortical processing stages, the cortical processing does not reduce the dimensionality of the signal anymore, it is probably not compact (Field, 1994). Instead of removing further redundancies, the visual cortex is more likely to extract and describe the remaining higher order redundancies, because they constitute the relevant information. Note, however, that orientation selectivity, the most prominent feature of the primary visual cortex, can be very well explained by the redundancy reduction principle (Dimitrov and Cowan, 1998).

It is highly challenging to find general principles or optimality criteria for the cortical representation. One coding strategy could be to extract hidden interdependencies between picture elements (like pixels) and obtain a transparent representation, in which the representing units or neurons are highly specialized to certain features or objects in the visual environment. Entities like these objects are likely to appear independently from each other, therefore this representation can be learned on the basis of statistical independence (see next section for further discussion). This is called factorial code.

The *decomposition* of the visual input into different submodalities and features—based on these higher-order correlations—essentially starts in the primary visual cortex (V1). Specialized groups of cells extract information about the different aspects of the visual scenes. Form, motion, or depth (and their combinations) at each location of the visual field are processed by highly interconnected parallel pathways. These pathways are specialized to the different modalities, but they strongly modulate each other via the extensive interconnectivity. Furthermore, neurons are specialized to features, such as orientation, corners or more compound patterns. The decomposition into different modalities and features results in an *overcomplete* representation in the cortex. The dimensionality of the representation in V1 increases by several factors compared to the retinal or geniculate representation: 260 million neurons process the feed-forward input from 2 million LGN fibers. The complexity of the extracted features increases from V1 towards higher visual areas (like MT), where one can find the prototype of "grandma cells", the face selective neurons. As a consequence, in higher areas fewer neurons represent a given sensory signal.

The cortical representation is a surreal mosaic of knowledge pieces at different complexity levels. The decomposition into "mosaic pieces" decreases the complexity of the neural representation and it is conducive to building new associations that may help to interpret the sensory signals. A pattern of light intensities is transformed into neural activity patterns, where the activity of the individual neurons or neuronal populations account for the presence of meaningful objects or complex features. This representation reflects semantic aspects of the input. Naturally, there is a need to combine or bind these little mosaic pieces into the coherent image we perceive. The anatomical structures or core networks that are specialized to certain features or submodalities are strongly interconnected. The binding of different features is established via recurrent interconnectivity. Due to this interconnectivity, perception of a coherent image emerges. Furthermore, visual cues from different submodalities can support the interpretation of others. (Several illusions arise from this effect, e.g., when the luminance gradient indicates an illusionary three-dimensional structure.)

Summarizing this brief introduction, the visual system *decomposes* and *integrates* the visual scenes in parallel. The integrative processes maintain the illusion of perceiving coherent images, while as a result of the decomposition principle, we are also able to analyze visual scenes at several levels of complexity: we can identify objects, colors, motion, forms, or faces. The neural mechanisms of the image decomposition and integration are well hidden by our own nervous system. Normally these mechanisms cannot be approached consciously. Introspection is usually a hopeless method for understanding them. However, in certain situations, when there is an obvious discrepancy between the actual reality as we know it and our perception, we can see behind the curtain. As our reference system is the real world and not our neurophysiological reality, we call these effects illusions. Illusions reveal some neural processing principles or connectivity structures that are otherwise well hidden from our "eyes". It is interesting to note that the visual system can be most easily tricked with the help of weird, artificial stimuli. This indicates that the visual system is specifically prepared (through evolution and learning) to interpret our natural environment (see also in section 1.2 and chapter 4).

Visual sciences are not even scratching the surface of the principles underlying the integrative processes involved in perception. However, the decomposition strategy of the visual system, is quite well described, although not too well understood. The representation in V1 and also in other early visual cortical areas is relatively transparent in the sense that activity of single cells correlates well with the presence of simple features in the visual stimulus. For instance, orientation selectivity, first described by Hubel and Wiesel (1962), is a prominent property of several cells in V1. Orientation selective cells respond strongly to edges or gratings with a certain orientation, but they remain inactive if the orthogonal orientation is presented. Even though the discovery of orientation selectivity was a revolutionary step to approach coding strategies in the neocortex, since then most of the research has been descriptive rather than interpretive. Substantial knowledge has been gained about the anatomical structures and neurophysiological mechanisms that are responsible for the generation of orientation selectivity (see chapter 2 and Das (1996); Vidyasagar et al. (1996); Sompolinsky and Shapley (1997) for reviews), but there have been very few studies aiming to understand the reason for the presence of orientation or other feature selectivity in the visual cortex in terms of signal processing. In some sense, the experimental description of the visual cortical function is well ahead of our theoretical understanding of it. Motivated by and based on the huge amount of available experimental data, several "bottom-up" computational models have been proposed. These models provide a better understanding of the neural "wetware" by deducing higher-order function from anatomical and neurophysiological observations. However, some of the basic questions are still open, including the origin of orientation selectivity. Therefore, we explored a computational model to highlight a new alternative hypothesis about the intracortical generation of orientation selectivity (chapters 2 and 3).

Certainly, research cannot continue without gaining some understanding of the principles of cortical representation and without providing a quantitative description of the neural responses in connection to the represented world. Recent experimental studies describing neural responses in V1 have made this need even clearer. These measurements demonstrated that the response of the orientation selective cells in V1 does not solely depend on the orientation of an edge within the "classical receptive field"[3], but it can be strongly *modulated* by patterns placed several degrees outside (e.g., Sillito et al., 1995; Zipser et al., 1996; Levitt and Lund, 1997; Polat et al., 1998). These

[3]The classical receptive field is the area from where neural firing can be evoked by localized visual stimulation.

reports, however, contradict each other at several points. The discrepancies are mainly[4] due to slight differences in the interpretation of the neural responses and the stimulus setup in the individual experiments. Stimulus setup is a sensitive issue because the parameter space explodes using compound stimuli assembled from gratings or bars with different contrast, orientation, or spatial frequency. Unfortunately, there is no general theory that could guide and motivate certain setups and systematic exploration of the neural response to the visual input. There is a strong need to point out "interesting" directions in this large parameter space because the number of available recordings is strongly limited by technical constraints.

One possible way to address the above questions and determine "interesting" directions in the stimulus space for the visual cortex is a careful and quantitative statistical or information theoretical (see next section) analysis of the recorded neural responses to statistically characterized stimulus sets. This approach has been successfully applied especially in the examination of the sensory systems of insects (see, e.g., Laughlin, 1994; Rieke et al., 1997), but interesting new research is conducted in mammalian cortex too (e.g., Richmond and Optican, 1990; Sugase et al., 1999). The other possible way is to reveal coding principles (such as the abovementioned factorial code) employed by the neocortex that may be optimal for sensory processing (for more detailed discussion see next section and, e.g., Barlow, 1961; Linsker, 1989; Olshausen and Field, 1996). Following the latter direction of research strategy, in the present thesis we deduce neurophysiological predictions starting from optimality requirements for the cortical encoding. From high order function we proceed towards the biological reality. This approach is referred to as "top-down" modeling. In close connection with our investigation into the origin of orientation selectivity (chapters 2 and 3), first we explore the role of feed-forward and recurrent lateral connections in obtaining an efficient cortical representation *in time* (chapter 4). Secondly, we study contrast adaptation (chapter 5) and we derive a learning rule for the transmitter release probability at the geniculo-cortical synapses. Both approaches are based on the principle of maximizing the *mutual information* between cortical input and output.

## 1.2 Information theory meets sensory processing?

In the following we shortly discuss why classical information theory (Shannon and Weaver, 1949) could provide a useful "toolbox" for investigating neural representation. In parallel, the basic terminology is introduced. More comprehensive reviews can be found in (e.g., Atick, 1992; Rieke et al., 1997).

The word "*information*" has a fairly complicated, multilayered meaning. What do we associate with the word "information"? Excitement, novelty, learning, structure, semantics... To be sure, it is scarcely a trivial task to give a mathematically exact and useful definition for "information" that also fits our common sense interpretation. Therefore the most widely used formal measure of information is not intended to be a formalization of our subjective concept of information. Shannon and Weaver (1949) introduced information theory for solving problems of telecommunication systems. This information measure is called Shannon information. Surprising as it may sound, information theory has recently proven to be a powerful tool of investigation into the neural code and the representation of sensory signals.

Let us put aside for a moment the promotion of Shannon information and imagine a phone

[4] Technical difficulties could also cause discrepancies in the results, but these are not in the scope of the current discussion.

cable. Beyond doubt, there are several differences between a phone cable and the central nervous system, but let us point out only one of them: the phone cable does not "care" whether we transmit the *Ulysses* by James Joyce, the Berlin phone book, or this thesis through it. Clearly, some human observers would find all of these three information streams absolutely boring or irrelevant, but some may not. Shannon and Weaver (1949) introduced an objective measure of information from the perspective of the phone cable. The measure ignores semantic aspects. Given a stochastic information source (denoted here as a random variable $S$) with the probability structure P(S) and the set of possible signals or alphabet $\mathcal{A}$, the *Shannon information* of a signal $s \in \mathcal{A}$ is

$$I(s) \equiv -\log_2 P(s) \,. \tag{1.1}$$

In the following discussion the term "signal" shall designate any type of representation of a message. This could be an image, text, sound, or neural activity pattern. Shannon information indicates the reduction of uncertainty by communicating message $s$. The unit of information according to the above definition is *bit* because the logarithm has a base of 2. For instance, telling the gender of somebody we have met has an information content of 1 bit because the number of possible people is reduced to the half (assuming that the male-female ratio in the set of the possible people is 1:1). Following our example, telling the name of the person has a high information content, assuming that each individual has a relatively low likelihood. Shannon information is the measure of unexpectedness. The less likely or more surprising an event or a message, the larger its information content.

The average information content of the possible signals is called *entropy*

$$H(S) \equiv -\sum_{s \in \mathcal{A}} P(s) \log_2 P(s) \,. \tag{1.2}$$

The entropy of a signal source is the number of bits that are on the average necessary to encode the signal. The signal transmission of the information is most efficient if it uses the same amount of bits on the average as the entropy of the signal. If more bits are used for the encoding than it is necessary, then the representation is redundant.

Entropy measures the average reduction in uncertainty by making one observation. If on the average one observation reduces the possible set of signals to the half, then its entropy is 1 bit. The entropy is maximal if every signal is equally likely. If some of the signals occur frequently, but others only rarely, the observer could expect the frequent event with a good chance even before making the observation. In this case the entropy is lower, and the efficacy of the communication between signal source and the observer is suboptimal.

Now consider the case in which the signal $s$ is a combination of $n$ symbols ($\vec{s} = s_1, \ldots, s_n$). Images, e.g., are composed of individual but usually not independent pixels. In this case the signal is a vector of the components, and its entropy satisfies

$$H(S) \leq \sum_{i=1}^{n} H(S_i) \,. \tag{1.3}$$

The equality holds if and only if the individual components are statistically independent. Correlations between the components (e.g., pixels) decrease the entropy and therefore the coding efficacy because by observing one pixel one could infer the value of the other.

To sum up, the signal source can transfer information with the highest rate to the observer if each signal is equally likely and the signal components are statistically independent from each other.

If the number of bits available for transmitting a signal stream (e.g., a phone conversion) is limited, then it is essential to obtain an optimal encoding of the given signal set (e.g., the human speech). The maximal entropy that can be theoretically reached in an information channel is called *channel capacity*

$$C \equiv \max_{P(\vec{s})}\big(H(S)\big) = \max_{P(s_i)}\left(\sum_{i=1}^{n} H(S_i)\right) = n\log_2 N\,, \tag{1.4}$$

where $N$ is the number of different symbols at each component of $\vec{s}$. Based on this, one could calculate, e.g., the capacity of 100 by 100 image matrixes if 256 different gray levels are available. Similarly the capacity of written text that consists of $N$ letters can be determined. Determining the capacity of a coding scheme is essential to characterize the efficiency of the information transfer. It turns out, e.g., that natural languages do not use the full capacity provided by the set of available letters. If the random variable is discrete, then the channel capacity is constrained by the available components and signals. Shannon information, and entropy can be extended for continuous random variables. In the continuous case, the capacity is constrained by the maximal available signal-to-noise ratio on the channel.

If the full channel capacity is not used, then the coding is redundant, less efficient. Given the channel capacity, the *redundancy* of a signal source $S$ is

$$R \equiv 1 - \frac{H(S)}{C}\,. \tag{1.5}$$

Redundancy gives a measure for the inefficacy of the signal transmission. One can distinguish between two types of redundancy sources by reformulating the above definition

$$R = \frac{1}{C}\underbrace{\left(C - \sum_{i=1}^{n} H(S_i)\right)}_{\text{first-order}} + \frac{1}{C}\underbrace{\left(\sum_{i=1}^{n} H(S_i) - H(S)\right)}_{\text{higher-order}}. \tag{1.6}$$

The first term increases if the different symbols in the available alphabet are used with unequal probability. This is referred to as first order redundancy. The second term increases if there are interdependencies between the signal components (e.g., between pixels). This is referred to as higher order redundancy. Natural languages are redundant both because of too frequent use of certain letters, like "e", and interdependencies between letters in a sequence (e.g., consonants are likely to be followed by vowels).

**What is the goal of neural coding?**

Having introduced the basic terms of information theory, let us return to our original problem of applying information theory in the context of neural representation. As it was emphasized before, Shannon information ignores semantic aspects of the signals. Instead, it considers a stochastic signal source with a given statistical structure $P(S)$ and an information channel. Transmitting information is problematic if the channel capacity is limited. This is called the "information bottleneck problem". Such limitation can arise from noise, constrained number of representational units and limited dynamic range of the activity (limited alphabet). Small subsystems of the nervous system could be also

considered as units that are "blind" to semantics of the signal, and their only task is to transmit information to each other in an effective manner (Atick, 1992). Based on this paradigm, Barlow (1961) proposed that the goal of sensory coding is to completely reduce the redundancy that is present in the stimulus. This is called *minimal redundancy* code.

Alternatively, Field has pointed out (Field, 1987; Field, 1994) that the goal of the cortical encoding is to minimize the statistical dependency between neurons (the second term in Eq. 1.6) in such way that the output entropy $H(S)$ nonetheless remains constant. This code is called *minimum entropy* or *factorial code*. If the neurons are statistically independent, decoding can be based on looking at the activity of the individual neurons, without considering complex interdependency patterns of firing.

Minimal entropy code alone does not reduce redundancy, and therefore it is not motivated by an information bottleneck as the minimal redundancy code. Minimal entropy or factorial code transfers the higher order interdependencies into first order redundancies. This means that the firing histograms of single neurons are redundant, neurons are more likely to fire with a certain frequency. In this code the entropy of the *single* components $H(S_i)$ is minimized (that is why it is called "minimal entropy code"). Several studies indicate that the orientation selective simple cells form a sparse and factorial representation of the natural world (Field, 1987; Olshausen and Field, 1996; Bell and Sejnowski, 1997; Olshausen and Millman, 2000). In other words, minimal entropy code is obtained in the primary visual cortex by the orientation selective cells.

Statistically independent components of natural scenes could be related to independent objects, causes or semantic units. Therefore, a factorial representation could also be advantageous for higher order cognitive function. Furthermore, after obtaining a factorial code, a subsequent simple gain control mechanism can map the neural output such that it obtains maximal entropy given a limited dynamic range of activity. If the interdependency between coding units is minimal, a smaller dynamic range is enough to gain a given amount of information about the signal. Combining factorial code with a proper gain control minimizes the redundancy in the neural code.

The observation that factorial code allows to reduce the dynamic range of the neurons while the information transfer is kept constant is used by Atick and Redlich (1990). They demonstrated that the receptive field shapes of retinal ganglion cells can be explained by the minimal redundancy principle. They suggest that redundancy can be reduced by reducing the channel capacity (the dynamic range of neural response) subject to the constraint that the information gained about the sensory input is equal to a given minimal required information. In other words, they propose that the neuronal channel capacity should not be wasted for encoding the input noise. The optimal solution is a receptive field, which decorrelates the sensory signal. Atick and Redlich (1990) have shown that considering only second-order correlations is enough to explain receptive field properties of the retinal ganglion cells. If output noise is also present, then the system increases correlations in the signal because correlations distinguish signal from noise (the noise was assumed to be uncorrelated on the different units).

#### Neural code and the (interactions with the) represented world

It is essential that the optimal coding strategy depends on the actual *a priori* distribution of the encoded stimulus. Natural images are not random, they have an inherent structure. They form a very small subset of all possible images, in other words, they are redundant (compared to what could be communicated by photon-beams). It has been argued that the neural representation is matched to this statistical structure (e.g. Field, 1987; Laughlin, 1981; Atick and Redlich, 1990; Atick and

Redlich, 1992; Field, 1994; van der Schaaf and van Hateren, 1996). If the visual system is specialized to our visual environment, understanding the structure of the visual world could guide us in the understanding of the visual system (Gibson, 1966). Redundancy in the neural code can be minimized by eliminating the redundancy in the sensory input. As first- and second-order correlations do not carry relevant information it is an efficient coding strategy to remove them. However, higher-order correlations make a visual scene, text, music *meaningful* for us. Extracting and describing these correlations, for example by obtaining a factorial code, is therefore a better strategy then removing them (Field, 1994).

In order to reduce or extract redundancies from natural stimuli, it is necessary to learn the environment's statistical structure. This continuously observed structure or redundancy becomes part of our *knowledge* about the world. Here, we use the word "knowledge" with a wider meaning that includes the hard-wired, genetically determined structures (obtained through the evolution) as well as the the knowledge accumulated from personal experiences and learning. Our expectations, our knowledge about the correlations in the outside world are crutches in the process of interpreting, learning and detecting novelty. Even though white noise has the highest information content in Shannon sense, it is meaningless for a human observer because our neural system is not able to cope with such high information rate.

We would like to remind the reader that the main assumption in the previous discussion was that neural subsystems need to communicate with each other in an efficient way, given a limited channel bandwidth. To reach this goal one strategy is to remove the redundancies from the original signal and obtain a compact code with minimal redundancy. Several studies indicate that subcortical systems do remove low-order redundancies. The coding strategy for cortical processing is likely to be different. It tries to transfer the complex interdependencies in the signal into simple first-order redundancies. As a consequence, single coding units represent independent objects that are likely to be meaningful. Note that once the higher order interdependencies are extracted, it is easy to reduce the first order redundancy in single neurons' activity. This can be done by a proper input-output mapping that matches the prior probability distribution of the represented object.

To sum up, we have shortly illustrated that considering the one-way chain of signal-channel-receiver, information theory can explain some important aspects of neural function in early sensory systems. However, this framework is likely to be too restricted to approach the high level cortical processing. It may be insufficient to analyze the high dimensional input space alone. Instead of considering the statistical structure solely of the signal (i.e. the outside world), one may investigate the structure of the *interactions* with the outside world. This extended framework could provide a description of coding in a semantic space that is defined based on *behavioral relevancy*.

Our research presented in this thesis focuses on two aspects of optimal coding in the primary visual cortex. Firstly, in the context of contrast adaptation we argue that the visual system adapts continuously to the slowly changing environment. We propose that the functional role of contrast adaptation is to eliminate the dependence of cortical activity on the root mean square (r.m.s.) contrast that soon after a change becomes a source of redundancy. In parallel, we investigate the low-level neural mechanism that could be responsible for "implementing" contrast adaptation. These bottom-up and top-down approaches meet at the learning rule for the transmitter release probability. Secondly, we investigate the short term dynamics of the recurrent competition and neural activity in the primary visual cortex in terms of information processing and in the context of orientation selectivity. We consider a free-viewing scenario where the environment is explored by fixating subsequently at different positions. This study is based on two key observations. *(i)* The coding strategy that maximizes information transfer depends on the signal-to-noise ratio (SNR) of the output. *(ii)* The SNR

changes with time within the duration of a fixation period because with increasing time window, the number of spikes available for encoding increases. It follows that the neural coding strategy should be modulated on a fast time scale to fit to the decreased noise level. We refer to this principle as "dynamic coding". In the context of orientation selectivity in V1, we suggest that it is optimal to decrease recurrent competition among different edge detectors tuned to different orientations after stimulus onset (after the beginning of a fixation period). This results in a complexity based, hierarchical feature extraction. The hierarchical levels are distributed in *time*. The model predicts that in the first phase of a fixation period, salient and typical features are processed. In the second phase, less typical, detailed structures are represented.

## 1.3 Structure of the thesis

Each chapter is self-contained with an extensive review of the previous work and an introduction to the discussed subject.

- In chapters 2 and 3, we explore a new hypothesis about the origin of orientation selectivity in V1. According to our "intracortical hypothesis" anisotropic recurrent projections provide *both* the initial orientation bias and its subsequent amplification. In chapter 2 we setup a rate model for an orientation hypercolumn and study the emergent receptive fields and the average response dynamics. Our study shows that the emerging response properties are similar to the response properties that are observed experimentally, hence the hypothesis of an intracortical generation of orientation bias and sharp orientation tuning is a sensible *alternative* to the notion of a feed-forward bias by convergent geniculo-cortical projection patterns that is subsequently sharpened by the recurrent connections.

- In chapter 3, the intracortical hypothesis is further explored. Using a statistical neural field approach we study the recurrent connectivity pattern and response dynamics in a greater biological detailness. We find that the anisotropy in the recurrent connections that is capable for generating the orientation bias does not need to be extreme strong, and therefore it may not be obviously reflected in the recurrent cortical patterns. Furthermore, we can account for the experimental observation that orientation tuning of the membrane potential sharpens gradually, while the spiking activity shows an immediate sharply tuned response.

- In chapter 4, the role of recurrent cortical amplification and its short-term dynamics is studied in terms of information processing. We propose that the recurrent cortical competition should decrease after stimulus onset to obtain optimal amount of information about the feed-forward signal in any increasing time window. The main motivation for the changing cortical competition comes from the observation that with increasing time window the signal-to-noise ratio in the cortical neural activity increases because more samples are available for representation and estimation. In the first phase of processing, only the most salient orientation is extracted by the strong recurrent competition. In the second, less competitive phase, the representation detailed structures and, therefore, multiple orientations becomes possible. Our information theoretic hypothesis is studied in an abstract model for a recurrent network. A more detailed model is also explored in order to give experimental predictions for the cortical response to multiple bars.

- In chapter 5, contrast adaptation, thus cortical dynamics on a longer time scale is studied. Firstly, we propose that a novel form of synaptic plasticity may be responsible for contrast adaptation. In contrast to the classical paradigm of modulating the synaptic strength that scales the *amplitude* of the synaptic transmission, we suggest that adaptation of the *dynamics* of synaptic transmission is the neural mechanism that explains all the available experimental data. In our model the synaptic dynamics is changed via the transmitter release probability. We further propose that switching between different modes of synaptic transmission has functional advantages in the context of contrast adaptation. Secondly, we derive an adaptation rule the transmitter release probability based on the hypothesis that contrast adaptation serves to achieve the most efficient cortical representation of the feed-forward input arriving from the lateral geniculate relay cells.

# Chapter 2

# Generating orientation selectivity intracortically—a rate model


## Abstract

We report results of numerical simulations for a model of generation of orientation selectivity in macaque striate cortex. In contrast to previous models, where the initial orientation bias is generated by convergent geniculate input to simple cells and subsequently sharpened by lateral circuits, our approach is based on anisotropic intracortical excitatory connections which provide *both* the initial orientation bias and its subsequent amplification. Our study shows that the emerging response properties are similar to the response properties that are observed experimentally, hence the hypothesis of an intracortical generation of orientation bias is a sensible *alternative* to the notion of an afferent bias by convergent geniculo-cortical projection patterns. In contrast to models based on an afferent orientation bias, however, the "intracortical hypothesis" predicts that orientation tuning gradually evolves from an initially nonoriented response and a complete loss of orientation tuning when the recurrent excitation is blocked, but new experiments must be designed to unambiguously decide between both hypotheses.[1]


## 2.1 Introduction

The emergence of orientation selectivity in the primary visual cortex of higher mammals has been one of the most active areas of research during the past decades (for recent reviews see Das, 1996; Sompolinsky and Shapley, 1997). Currently favored models assume convergent thalamic feed-forward projections from elongated regions of the thalamic visual field representation which are complemented by strong intracortical recurrent connections. According to those ideas, the afferent projection generates an orientation bias which is subsequently amplified by intracortical circuits. While the role of intracolumnar recurrent excitation and inhibition in the generation of tuned responses has been clarified at least to some extent, the origin of the initial orientation bias, suggested to be a property of the geniculo-cortical projection, is debatable.

[1] This chapter is based on (Adorján, Levitt, Lund and Obermayer, 1999).

The idea of an afferent orientation bias goes back to Hubel and Wiesel (1962) who proposed that simple cells in cat striate cortex are orientation selective because they receive segregated ON- and OFF-input from appropriately elongated areas of the LGN retinotopic map. According to their hypothesis, the axis of elongation of the afferent projection determines orientation preference while the aspect ratio of the receptive-field determines the specificity of the response. Estimates of the aspect ratio of the cortical cells' anatomical receptive-fields, however, are in conflict with the sharp tuning of simple cells (Chapman et al., 1991; Pei et al., 1994; Reid and Alonso, 1995; Sompolinsky and Shapley, 1997) and - as a purely feed-forward model - Hubel and Wiesel's hypothesis could not account for response properties such as the contrast invariance of the tuning width (Sclar and Freeman, 1982) or the sensitivity of the orientation tuning curve to changes in the strength of lateral inhibition (Tsumoto et al., 1979; Sillito et al., 1980; Eysel et al., 1990; Sato et al., 1996; Crook et al., 1997; Crook, Kisvárday and Eysel, 1998).

The fact that the lateral intracortical excitatory contribution to a cortical cell's synaptic input in layer 4 is much bigger ($80\% - 95\%$) than the afferent LGN input (see, e.g. Peters et al., 1994), the results of experiments blocking inhibition, and evidence for cross-orientation suppression from physiological studies (Hata et al., 1988; Bonds, 1989; DeAngelis et al., 1992) motivated the extension of Hubel and Wiesel's hypothesis to include lateral inhibition between cortical cells. In particular cross-orientation inhibition was thought to be a good candidate for sharpening orientation tuning, and model studies (see, e.g. Wehmeier et al., 1989; Wörgötter and Koch, 1991; Sabatini, 1996) demonstrated that an initially weak orientation bias can indeed evolve into a sharply tuned response by these kinds of interactions. A number of studies have reported finding cross orientation inhibition (Hata et al., 1988; Bonds, 1989; Douglas et al., 1991; DeAngelis et al., 1992; Pei et al., 1994) and blockade of cross orientation inhibition reduces orientation selectivity (Eysel et al., 1990; Wörgötter and Eysel, 1991; Crook et al., 1997; Crook, Kisvárday and Eysel, 1998). On the other hand, Ferster (1986) reported the lack of any cross-orientation hyperpolarization, and showed that IPSPs evoked by visual stimulation were actually strongest at the preferred orientation. Furthermore, the absence of cross-orientation shunting inhibition was also shown (Douglas et al., 1988; Ferster and Jagadeesh, 1992). When Berman et al. (1992) suggested that the measured small changes in membrane conductance and the measured small hyperpolarization are not sufficient to cancel strong monosynaptic afferent inputs and Nelson et al. (1994) reported that blocking inhibition within a single cell does not affect its orientation tuning, it became clear that an essential ingredient was missing from these models of generation orientation specificity.

These findings led to the current set of hypotheses concerning the origin of orientation selectivity, which include strong recurrent lateral excitation in addition to iso- and cross-orientation inhibition (Douglas and Martin, 1991; Ben-Yishai et al., 1995; Douglas et al., 1995; Somers et al., 1995; Mundel et al., 1997; Carandini and Ringach, 1997). According to these hypotheses, recurrent iso-orientation excitation selectively amplifies the weak afferent signal while iso-orientation and cross-orientation inhibition are required for the control of the amplification and for the sharpening of the initial orientation bias respectively. These models were, finally, able to reconcile a large body of data and lead to a consensus about the role of local lateral interactions in orientation tuning.

Returning to the evidence for an afferent vs. a cortical origin of the initial orientation *bias*, experimental results for cat and ferret (Chapman et al., 1991; Reid and Alonso, 1995; Ferster et al., 1996) seem to support the idea of thalamic feed-forward generation of an orientation bias. Pei et al. (1994), Volgushev et al. (1995), and Ringach et al. (1997*a*), on the other hand, describe many cells with an initially nonspecific response which evolves to a sharply orientation tuned response after approximately 10 msec, contradicting observations of immediate, strongly tuned afferent EPSPs

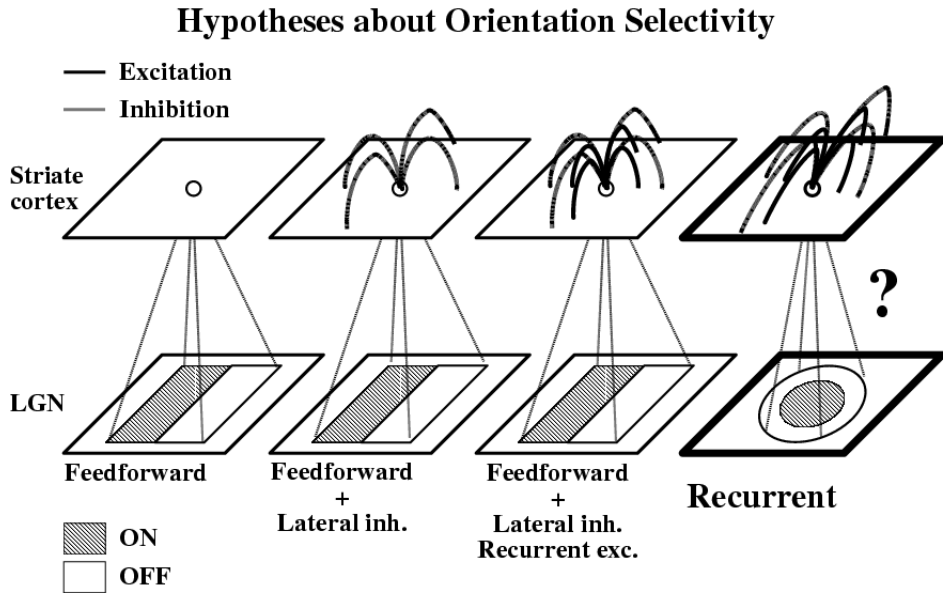


Figure 2.1: Hypotheses about orientation selectivity: *(i)* convergence of the feed-forward connections from an elongated region in the LGN is solely responsible for orientation selectivity; *(ii)* recurrent inhibition sharpens the feed-forward orientation bias; *(iii)* cortical amplifier models, recurrent excitation and inhibition sharpens the feed-forward orientation bias; *(iv)* the "intracortical hypothesis", the initial orientation bias is generated and sharpened by the anisotropic recurrent connections.

(Ferster et al., 1996). Also, inhibitory blockade experiments indicate circular rather than elongated excitatory receptive-fields (Tsumoto et al., 1979; Sillito et al., 1980). Indirect evidence for a cortical component of the orientation bias, also comes from recent developmental studies (Kim and Bonhoeffer, 1994; Bonhoeffer and Goedecke, 1996). Reverse suture experiments performed in kittens and young ferrets left the visual cortical orientation map unchanged, even when the eyes never had common visual experience. If monocular deprivation results in a complete reorganization of the afferent projection as has been reported by (Antonini and Stryker, 1993) then it is difficult to explain (but see Wolf et al. (1996) for an attempt) that "connecting" or "reconnecting" of fibers after reverse suture can lead to the restoration of the same biases in the afferent projections and to virtually identical cortical orientation maps.

Given the abovementioned contradictory evidence, the fact that orientation selectivity in primates has not yet been extensively addressed, and the fact that nearly all previous model studies consider only an afferent origin of the orientation bias motivated us to explore an *alternative hypothesis*, namely that *both* the initial orientation bias and the subsequent amplification are generated by the same specific lateral excitatory recurrent connections. Thus we focus on the mechanisms underlying the initial symmetry breaking in the orientation domain which may—in principle—be small (Ben-Yishai et al., 1995) but which in reality has to be large enough to robustly overcome noise. The goal of this study is twofold. Firstly, the predictions of the "intracortical hypothesis" are explored and the receptive-field properties—orientation selectivity, spatial frequency tuning, and

spatial receptive-fields—are derived. Secondly, we look for experiments to test our "intracortical hypothesis". The goal of our study is then to show that—given the current evidence—intracortical generation of orientation selectivity is a viable alternative to the "afferent" hypothesis and needs to be explored seriously.

In the following section we describe the structure of our model and provide the model parameters. Section 2.3 contains the results of our numerical simulations with respect to the contrast dependence of orientation tuning, the dynamics of the response, the role of lateral excitation and inhibition in orientation selectivity, and with respect to the spatial frequency tuning and spatial structure of the receptive-fields. Section 2.4 contains a critical discussion of model assumptions, a comparison of model predictions with experimental data, a comparison with other models of orientation selectivity, in particular with models which assume an afferent orientation bias, and a discussion of some experimentally testable model predictions. The model is based on the tuning properties of simple cells in layer 4C$\alpha$ of macaque striate cortex, for which an anatomical substrate for the generation of orientation selectivity has recently been suggested (Yoshioka et al., 1994). A preliminary study on the "intracortical hypothesis" was published in (Bauer et al., 1997).

## 2.2 Methods

### 2.2.1 Overview

As a basis of our modeling study we chose orientation selective cells in layer 4C$\alpha$ of the striate cortex in the macaque monkey. This seems an appropriate choice for three reasons. (i) Orientation selective cells are found for the first time in mid and upper layer 4C (Blasdel and Fitzpatrick, 1984; Hawken and Parker, 1984) and they coincide with the emergence of lateral axon projections from the excitatory spiny stellate neurons reaching up to 500 to 1500 $\mu$m along one axis from the cell bodies (Lund, 1987; Anderson et al., 1993; Yoshioka et al., 1994), which may actually serve as an anatomical substrate for generation of orientation selectivity. (ii) A recent report suggests that orientation tuning for some cells in primate layer 4C$\alpha$ is not established immediately after response onset but gradually develops after an initial nonoriented response within the first 10-15 msec (Ringach et al., 1997*a*, and personal communication). This phenomenon has also been observed in cat striate cortex (Dinse et al., 1991). In contrast Celebrini et al. (1993) reports immediately tuned cortical response, but a more detailed modeling study (Adorján et al., 1998) gave a possible solution for this contradiction. (iii) No attempt has yet been made to explain the emergence of orientation selectivity in macaque striate cortex, although its architecture differs in several important ways from the primary visual areas of cats and ferrets. It is possible that the mechanisms underlying the emergence of receptive-field properties differ between species.

Our computational model consists of three layers: the visual field layer, the magnocellular layer of the LGN, and a cortical layer which corresponds to the upper region of 4C$\alpha$ of primary visual cortex (Fig. 2.2). We chose to consider only monocular ON-cells, and have neglected ON-OFF interactions to simplify our model. We have done so because it has already been demonstrated that blocking the ON pathway does not alter orientation or direction selectivity of V1 cells (Schiller, 1982; Sherk and Horton, 1984). Furthermore, cross-correlation studies (Tanaka, 1983; Reid and Alonso, 1995) provide strong evidence that inputs from the LGN to the individual subfields of V1 simple cells are essentially segregated to the same type (i.e. ON to ON and OFF to OFF). Thus, ON-OFF interactions seem unlikely to be required for generation of orientation selectivity. Parameters

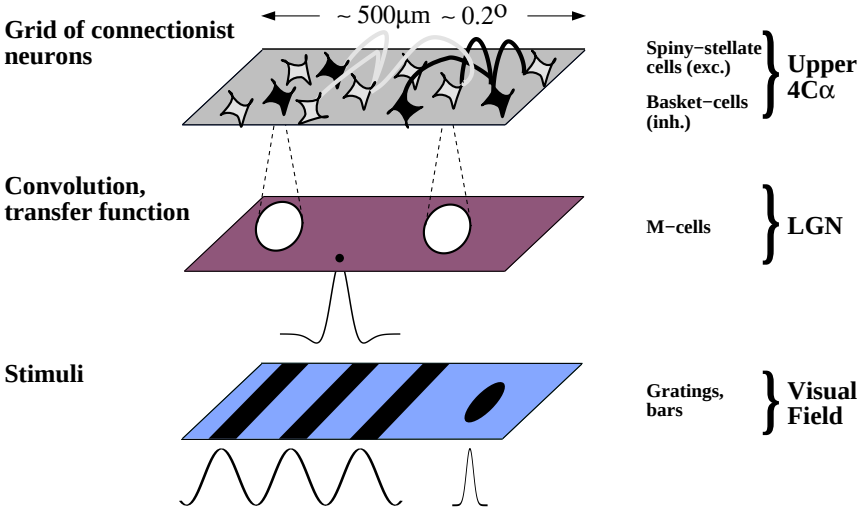


Figure 2.2: The structure of the connectionist model for generating orientation selectivity intracortically.

are taken from measurements at 5° eccentricity in the visual field representation. To simplify notation we will identify model neurons of the same kind by the location of their receptive-field center in visual space, and not by their anatomical locations in each layer.

### 2.2.2 Stimuli

Stimuli are stationary spots or gratings which are presented to the visual field layer. The luminance values $l_s$ for a spot stimulus are given by

$$l_s(\vec{x}) = 1 + c \cdot \exp(-(x_1 - u_1)^2/s_1^2) \exp(-(x_2 - u_2)^2/s_2^2)\,, \tag{2.1}$$

where $\vec{u} = (u_1, u_2)$ is the position of the spot's center in visual field, $c$ its Weber contrast ($(L_{\max} - L_{\min})/L_{\min}$), and $\vec{x} = (x_1, x_2)$ are visual field coordinates. A stationary sinusoidal grating $l_g$ is given by

$$l_g(\vec{x}) = 1 + c \cdot \cos(2\pi f \cdot d(\vec{x}, \alpha))\,, \tag{2.2}$$

where $c \in [0..1]$ denotes Michelson contrast ($(L_{\max} - L_{\min})/(L_{\max} + L_{\min})$), $\alpha$ the orientation, $f$ the spatial frequency, and $d(\vec{x}, \alpha) = x_1 \sin\alpha - x_2 \cos\alpha$. If not mentioned otherwise, all the simulations with sinusoidal gratings were made at the optimal spatial frequency of the geniculate M-cells.

### 2.2.3 The magnocellular layer

The receptive-field profiles $S$ of the geniculate M-cells are described by a Difference of Gaussians (DoG) model (Rodieck, 1965) and the afferent input $R(\vec{u})$ to a geniculate M-cell at location $\vec{u}$ in the magnocellular layer is given by the convolution of the stimulus $l_{s,g}$ in the visual field layer

with the receptive-field profiles $S$. Parameters of the receptive-field profiles were taken from (Spear et al., 1994) and were corrected for 5° eccentricity (see Appendix A.1). They are listed in Table 2.1.

| **Properties of geniculate M-cells (Spear et al., 1994)** | |
|---|---|
| Peak center sensitivity $k_c$ in [inp[a] (% c)$^{-1}$ deg$^{-2}$] | 2077.64 |
| Center radius $r_c$ in [deg] | 0.103 |
| Peak surround sensitivity $k_s$ in [inp (% c)$^{-1}$ deg$^{-2}$] | 14.75 |
| Surround radius $r_s$ in [deg] | 1.16 |
| Optimal spatial frequency $f_{\text{opt}}$ in [cycl/deg] | 0.59 |
| Maximal response $M_{\max}$ in [spikes sec$^{-1}$] | 43.00 |
| Contrast gain $G$ in [spikes sec$^{-1}$ (% c)$^{-1}$] | 1.82 |
| Activity gain $1/b$ [inp$^{-1}$] | 0.064 |
| Contrast threshold $c_{\text{M}}^{\min}$ in [% ] | 1.36 |
| Activity threshold $T_{\text{M}}$ [inp] | 7.77 |
| **Geniculo-cortical connectivity** | |
| Radius of the geniculate axonal arbor $r_{\text{ax}}$ in [$\mu$m] (Freund et al., 1989; Blasdel and Lund, 1983) | 400 |
| Radius of the dendritic arbor of cortical cells $r_{\text{dend}}$ in [$\mu$m] (Lund, 1980) | 100 |
| **Architecture of upper** $4C\alpha$ | |
| Magnification factor (Hubel and Wiesel, 1974) in [deg/mm] | 0.4 |
| Number ratio of excitatory to inhibitory cells (Lund, 1987) | 8:2 |
| Number ratio of excitatory to inhibitory synapses (Beaulieu et al., 1992) | 83:17 |
| Number ratio of afferent to lateral excitatory synapses (Peters et al., 1994) | 6:94 |
| Number ratio of synapses terminating on excitatory to synapses terminating on inhibitory cells (Freund et al., 1989; Anderson et al., 1994) | 9:1 |
| Activity gain $\beta$ | 0.13 |
| Activity threshold $T_{\text{C}}$ | 0.003 |
| Excitatory connection strength $W_{\text{C}}(e,\{e,i\})$ | 3.73 |
| Inhibitory connection strength $W_{\text{C}}(i,\{e,i\})$ | -24.5 |
| Specificity of excitatory connectivity (see Fig. 1b) in $[\frac{\%}{\text{deg}}]$ | -1.6 |
| Specificity of inhibitory connectivity (see Fig. 1b) in $[\frac{\%}{\text{deg}}]$ | -0.06 |

[a] Arbitrary scalable unit indicating the afferent input on a postsynaptic geniculate M-cell

Table 2.1: Summary of model parameters used in this study.

Because the emergence of the orientation selective response is much faster than the temporal modulation of the usually used stimuli, the output $O$ of the magnocellular layer is assumed to be

instantaneous, and it is calculated via a transfer function $g_{\mathrm{M}}$,

$$O = g_{\mathrm{M}}(R) = \begin{cases} 0 & \text{if } R \leq T_{\mathrm{M}} \\ \frac{R - T_{\mathrm{M}}}{R - T_{\mathrm{M}} + b} & \text{otherwise} \end{cases}, \tag{2.3}$$

whose parameters $T_{\mathrm{M}}$ and $b$ were determined according to Appendix A.2.

### 2.2.4 The geniculo-cortical connectivity

The total afferent input $A$ into a cortical cell is given by the convolution of the outputs $O$ of the geniculate M-cells with a circularly symmetric cone shaped weight kernel with radius $r_{\mathrm{A}}$,

$$W_A(d) = \begin{cases} \frac{r_A - d}{r_A} \cdot \frac{3N_A}{r_A^2 \pi} & \text{if } d \leq r_{\mathrm{A}} \\ 0 & \text{otherwise} \end{cases} , \tag{2.4}$$

whose integral is equal to the total number $N_{\mathrm{A}}$ of afferent synapses[2], and $d$ is the distance from the cell's receptive-field center in the visual field coordinate system. $r_{\mathrm{A}}$ is given by the radius of the region of overlap between the axonal arbor of a geniculate M-cell ($r_{\mathrm{ax}} \approx 400\mu m$, (Freund et al., 1989; Blasdel and Lund, 1983)) and the dendritic arbor of a cortical spiny stellate cell ($r_{\mathrm{dend}} \approx 100\mu m$, (Lund, 1980)). Because $r_{\mathrm{ax}} >> r_{\mathrm{dend}}$, we obtain $r_{\mathrm{A}} = r_{\mathrm{ax}} - r_{\mathrm{dend}} = 300\mu m$ in cortical coordinates and — given a cortical magnification factor of $0.4°/\mathrm{mm}$ — $r_{\mathrm{A}} = 0.12°$ in visual field coordinates. Note that there is no afferent orientation bias to an individual cortical cell because its receptive-field is circular symmetric.

### 2.2.5 The cortical layer

Our model of a cortical orientation cycle (orientation hypercolumn) consists of 300 neurons of which 80% are excitatory and 20% are inhibitory (Fitzpatrick et al., 1987; Lund, 1987; Beaulieu et al., 1992). The diameter of the hypercolumn is approximately $500\mu m$ which leads to a maximum distance between receptive-field centers in visual space of approximately 0.2° for cells located in the same orientation hypercolumn. The diameter of a cortical cell's receptive-field is approximately $0.2° - 0.3°$, thus receptive-fields of cells within one orientation hypercolumn overlap heavily.

For the purpose of the model we assume that cells whose receptive-fields lie along one axis in visual space belong to one "orientation column" (Fig. 2.3a) and are laterally connected by strong excitatory interactions. Inhibitory connections follow the same tendency in a less specific way. The strength of interaction between units from different orientation columns falls off with the angle between their axes (Fig. 2.3b). As will be seen later, the orientation of the axis in Fig. 2.3a determines the preferred orientation of its cells, and the interactions shown in Fig. 2.3b then correspond to iso-orientation excitation, iso-orientation inhibition, and weak cross-orientation inhibition. This choice of cortical connectivity in the orientation domain is reasonable given the physiological experiments which indicate highly specific iso-orientation excitation as well as strong iso-orientation inhibition (Ferster, 1986; Douglas et al., 1991), and the experiments showing the importance of inhibition in the emergence of the orientation selective response (Tsumoto et al., 1979; Sillito et al., 1980; Sato et al., 1996) or the presence of lateral inhibition arriving from oblique or orthogonally oriented cells (Hata et al., 1988; Bonds, 1989; Eysel et al., 1990; Wörgötter and Eysel, 1991; DeAngelis et al., 1992; Pei et al., 1994; Crook et al., 1997; Crook, Kisvárday and Eysel, 1998). We will show in section 2.3 that the model with this set of parameters fits to available experimental data and that these results are robust to changes in the specific choice of these parameters (cf., Figs. 2.7, 2.8). The overall distribution of different types of synapses targeting different cells is listed in Table 2.1. Excitatory connections originate either from the magnocellular LGN layer or from other cortical

[2] 6 %of the total number of synapses, which was arbitrarily set to 1000.

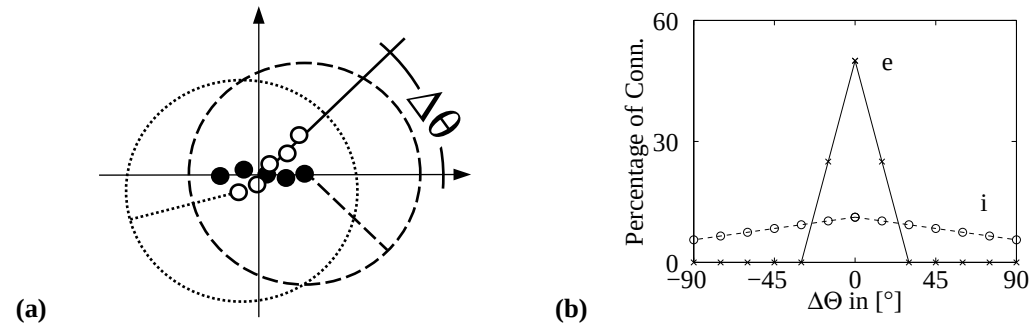


Figure 2.3: **(a)** Cartoon of an orientation hypercolumn. Filled and empty circles denote the centers of receptive-fields in visual space of cells from two orientation columns with a difference $\Delta\Theta$ in preferred orientation. The maximum distance between field centers is $\approx 0.2°$ and corresponds to the hypercolumn diameter at 5° eccentricity. receptive-field (dotted and dashed circles) diameters are approximately 0.3°. **(b)** Percentage $P(p,\Delta\Theta)$ (Eq. 2.8) of excitatory, e $\rightarrow$ e and e $\rightarrow$ i, (solid line) and inhibitory, i $\rightarrow$ e and i $\rightarrow$ i, (dotted line) lateral connections as a function of the difference $\Delta\Theta$ of orientation preference, as used in most of the numerical simulations. The particular choice of the lateral connectivity pattern was motivated by the available experimental data (see text) as well as by the numerical simulations which are shown later. The synaptic load is calculated by multiplying the number of synapses at the given difference in the preferred orientations $\Delta\Theta$ with the weight of a single synapse making contact with an excitatory or an inhibitory neuron. A typical set of parameters is: synaptic weights $W_C$: e $\rightarrow$ {e, i}=3.73, i $\rightarrow$ {e, i} = -24.5, slope of the percentage of the lateral excitatory and inhibitory connections $\frac{P(p,0)-P(p,15°)}{15°}$ as a function of the difference in preferred orientation: e $\rightarrow$ {e, i} -1.6 $\frac{\%}{\text{deg}}$ and i $\rightarrow$ {e, i} -0.06 $\frac{\%}{\text{deg}}$.

cells, while inhibitory synapses are made only between cortical cells. All three types of connections may terminate on excitatory or inhibitory cells.

Cortical neurons are modeled as continuous-valued units, whose state $m$ is interpreted as a "membrane potential". Let $\theta$ denote the preferred orientation of a cell as given by the orientation of the axes in Fig. 2.3a, index $i$ the position of its receptive-field center along each axis, $p \in \{\text{exc},\text{inh}\}$ the type of the cell, and $t$ the time. Then we obtain for the membrane potential $m$:

$$\frac{d}{dt} m^{(\text{p})}_{\Theta,\text{i}}(t) = -m^{(\text{p})}_{\Theta,\text{i}}(t) + I\,, \tag{2.5}$$

where the synaptic input

$$I = A + L \tag{2.6}$$

is the sum of the afferent input $A$ (section 2.2.4), and the lateral input

$$L(\Theta,q,t) = \sum_{p,j,\Theta'} N_{\text{C}}(p,q,|\Theta-\Theta'|)\, W_{\text{C}}(p,q)\, g_{\text{C}}(m(\Theta',p,j,t)). \tag{2.7}$$

$N_C(p,q,|\Theta-\Theta'|)$ is the number of synapses from a cortical cell of type $p$ targeting a cortical cell of type $q$ whose difference of preferred orientation is $|\Theta-\Theta'|$. $W_C(p,q)$ is the strength of a single connection between cells of type $p$ and type $q$. The percentage of the excitatory or inhibitory connections as a function of difference in preferred orientations $\Delta\Theta$ (Fig. 1b) is

$$P(p,\Delta\Theta)=\frac{N_C(p,\{e,i\},\Delta\Theta)}{\sum_{\Delta\Theta}N_C(p,\{e,i\},\Delta\Theta)}. \tag{2.8}$$

The cortical transfer function

$$g_C(x)=\begin{cases}0 & \text{if } x\le T_C\\ \beta(x-T_C) & \text{if } T_C<x<T_C+\beta^{-1}\\ 1 & \text{otherwise}\end{cases} \tag{2.9}$$

describes the transformation between the membrane potential and the output firing frequency. The strength $N_C W_C \beta$ of the recurrent amplification is being changed in section 2.3.2 by varying the strength of the different synapses $W_C$ (cf., Figs. 2.7, 2.9) at a fixed number of connections $N_C$ (see Table 2.1) and a fixed $\beta$ ($\beta=0.13$)[3]. The threshold $T_C$ was set to the geniculate input which a cortical cell in upper 4C$\alpha$ receives at threshold contrast ($c_C^{\min}=2\%$) (Hawken and Parker, 1984), $T_C=0.003$. The maximal activity of the cortical cells is 1.

### 2.2.6 Implementation

The model was implemented in C on a standard Unix workstation. The afferent input to a single cortical cell was integrated in space using the *extended Simpson's rule* (Press et al., 1994). The system of ordinary differential equations, Eq. 2.5, was integrated using *fourth-order Runge-Kutta* method (Press et al., 1994) with a time step of 0.5 (in arbitrary unit).

## 2.3 Results

### 2.3.1 Orientation bias and orientation tuning

In our model, all cortical cells receive geniculate input from a circular symmetric region of the magnocellular layer, hence the geniculate input to individual cortical cells does not provide an orientation bias. The orientation bias is generated by the anisotropic lateral interactions. The excitatory coupling is strongest between cells whose fields are located on a particular axis in the visual field (cf. Fig. 2.3) such that the sum of the total geniculate input arriving to all such coupled cells depends on the orientation of the stimulus (*columnar orientation bias*). In the model these strongly coupled cells have the same preferred orientation and form one "orientation column". The orientation bias of the summed geniculate input to such an orientation column is determined by the orientation of the axis on which the cells' receptive-fields lie. This *columnar orientation bias* is sharpened by the same recurrent connections, thus the generation of the initial orientation bias and the sharpening of the orientation tuning are inseparable processes.

[3] $\beta$ was chosen small enough such that the steady state activity of the network assuming a transfer function with no saturation is always less than 1 for all possible values of I (Eq. 2.6)

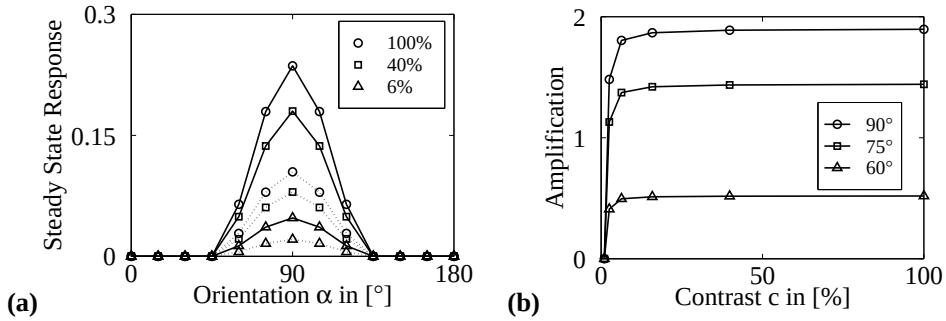


Figure 2.4: **(a)** Orientation tuning curves of excitatory (solid line) and inhibitory (dotted line) cells for three different values of stimulus contrast. The tuning width (half width at half height) is approximately 23° for both cell types and is independent of contrast. **(b)** Cortical amplification factor (steady state activity divided by the afferent geniculate input) as a function of contrast for three different stimulus orientations. Stimuli were gratings with optimal spatial frequency. Parameters were taken from Table 2.1 and Fig. 2.3.

Fig. 2.4a shows orientation tuning curves of a cortical cell for a grating stimulus of optimal spatial frequency but varying contrast. Tuning width is independent of contrast (cf. Sclar and Freeman, 1982) and is approximately equal for excitatory and inhibitory cells, as long as the specificity of the excitatory (e $\rightarrow$ e and e $\rightarrow$ i) and inhibitory (i $\rightarrow$ e and i $\rightarrow$ i) connections in the orientation domain is independent of the type of the target cell. Tuning widths (half width at half height) are approximately 23° for the parameters chosen but may vary with the connection specificity and the strength of the lateral interactions (see section 2.3.2). If the recurrent excitation is strong enough, the tuning width is to a large extent determined by the specificity of the lateral interactions. The tuning width no longer depends on the strength of the orientation bias, i.e. , the network operates in its "marginal phase" (Ben-Yishai et al., 1995). If the strength of the lateral excitation falls below a critical value, the tuning width becomes bias dependent (cf. Fig. 2.9).

Fig. 2.4b shows that the cortical amplification factor, defined as the steady state cortical response divided by the afferent geniculate input, remains constant with respect to stimulus contrast for strong enough recurrent excitation. This result indicates that a recurrent network exhibits either contrast invariant orientation tuning or saturation in the contrast response function, but not both if linear summation of the synaptic inputs and a piecewise linear transfer function is assumed. This result has been confirmed analytically by Bartsch et al. (1997). Hence additional mechanisms have to be invoked to explain the saturation of the contrast response curves, a finding which contradicts claims put forward in previous modeling studies (Todorov, Siapas and Somers, 1997).

### 2.3.2 The role of lateral inhibition

A large number of experiments in the past have suggested that lateral inhibition plays at least two roles in the emergence of orientation selective cells: it controls runaway excitation and it leads to the sharpening of the orientation tuning.

Blocking inhibition on a large cortical site with the $GABA_A$ antagonist bicuculline broadens

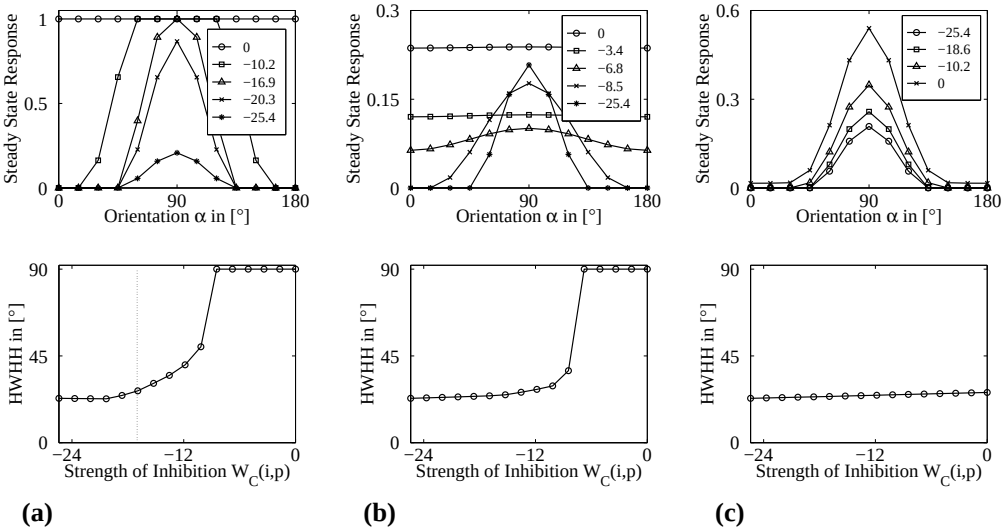


Figure 2.5: Orientation tuning curves and half width at half height as a function of **(a)** the strength $W_C$ of all inhibitory connections, **(b)** the strength $W_C$ of inhibitory connections for $\Delta\Theta > 30°$, with all other $W_C$ set to -24.5, and **(c)** the strength $W_C$ of inhibitory connections targeting on a single cell, with all other $W_C$ set to -24.5. The dotted vertical line in (a, bottom) indicates the strength above which the cortical response begins to saturate. The tuning curves (top) are plotted at five different parameter values, as indicated in the insets. Stimuli were gratings with optimal spatial frequency and $c = 60\%$ contrast. Parameters were taken from Table 2.1 and Fig. 2.3.

or diminishes the orientation selective cortical response, depending on its concentration (Tsumoto et al., 1979; Sillito et al., 1980; Sato et al., 1996). We simulated the effect of these experiments by reducing the strength $W_C$ of all inhibitory connections. The model then predicts a broadening of the tuning curve if the strength of inhibition is reduced to 70%or less (Fig. 2.5a (top)), but in contrast to experimental findings, the broadening of the tuning curve is mainly due to the fact that cortical cells saturate because they are driven into the flat region of their transfer function. Since real cortical cells saturate way below their maximum firing rate - with (Sillito et al., 1980) and without (see, e.g. Ohzawa et al., 1985) the presence of bicuculline - there must be additional mechanisms for the control of runaway excitation.

Inactivation of small cortical sites by micro-iontophoresis of GABA broadens orientation tuning when inactivation and recording sites have different preferred orientation (Eysel et al., 1990; Crook et al., 1997; Crook, Kisvárday and Eysel, 1998). To emulate the local inactivation experiments we changed the strength $W_C$ of inhibition between cells whose difference in preferred orientation was larger than 30°. The strength of inhibition is changed sequentially corresponding to the different levels of inhibitory blockade in the real experimental setup. The numerical simulations showed that the selective blockade of cross-orientation inhibition leads to an increased response at the null-orientation and/or to a decreased response at the optimal orientation. The tuning curve broadens significantly if the strength of inhibition is reduced to 30%or less (Fig. 2.5b), and activ-

ity now remains below the saturation level of the cortical cells. Iso-orientation inhibition, which remains active, is still sufficient to control runaway excitation.

Intracellular blockade of inhibition elevates the activity at every orientation by the same amount, but does not affect orientation tuning at simple cells (Nelson et al., 1994). We simulated this experiment by changing the strength $W_C$ of the inhibitory connections afferent to a single excitatory cell. The simulation results (Fig. 2.5c) indeed show a small elevation of the activity level, but the tuning width remained constant, because the cell is driven by sharply tuned recurrent excitation from the unaffected cells.

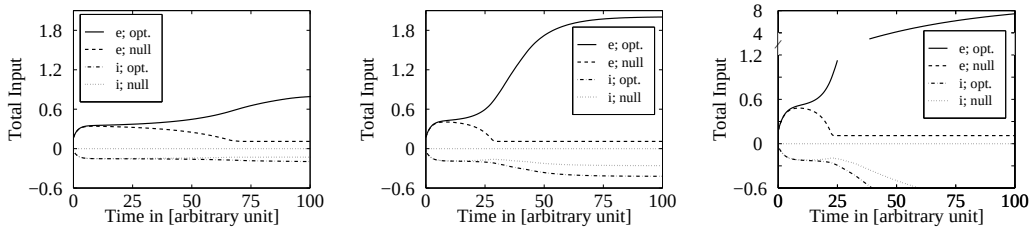


Figure 2.6: Total excitatory and total inhibitory input to a simulated cortical cell as a function of time for grating stimuli of null and optimal stimulus orientation and for different strengths of the excitatory connections: $W_C = 3.4$ (left); $W_C = 3.73$ (middle); $W_C = 3.9$ (right). Gratings had optimal spatial frequency and $c = 60\%$ contrast. Parameters were taken from Table 2.1 and Fig. 2.3.

Fig. 2.6 center shows the time course of the total excitatory and the total inhibitory input to a cortical cell for optimal and for null stimuli at the recurrent excitatory strength $W_C = 3.73$ generally used in our simulations. A fast unspecific rise of excitation is followed by the onset of inhibition. As inhibition rises, the activity of excitatory cells tuned to the null orientation decays while the activity of cells tuned to the optimal orientation grows. Finally, a tuned response emerges. Thus the model predicts a non-oriented initial response and a gradual sharpening of orientation tuning, similar to what has been reported by Pei et al. (1994) and Volgushev et al. (1995) for cat area 17, and for macaque V1, layer 4C$\alpha$ (Ringach et al., 1997*a*, and personal communication). This prediction differs from the predictions of models assuming a Hubel and Wiesel style afferent orientation bias which provides tuned input immediately (cf. Somers et al., 1995). Figs. 2.6 left and right show how the time-course of excitation and inhibition changes if the strength of the lateral excitatory connections decreased from $W_C = 3.73$ to $W_C = 3.4$ or increased to $W_C = 3.9$. If the lateral excitation is weak (Fig. 2.6 left) then the afferent nonspecific input becomes dominant and it invokes long nonspecific response. The emergence of the tuned response is delayed and the tuning width decreases. In contrast, at stronger lateral excitation (Fig. 2.6 right) the initial nonspecific response is very short. Tuning is established faster, but because the steady state activity is higher, the time to reach the steady state increases, as the tuning becomes sharper.

Since the strengths $W_C$ and the local projection pattern of the excitatory and inhibitory connections are free parameters, we explored the sharpness of orientation tuning (half width at half height) and the maximal steady state activity as a function of the strength of the excitatory and inhibitory connections (Fig. 2.7) and for different lateral connectivity schemes (Fig. 2.8). $N_C W_C \beta$ is approximately the slope of the postsynaptic membrane potential as a function of the presynaptic firing rate

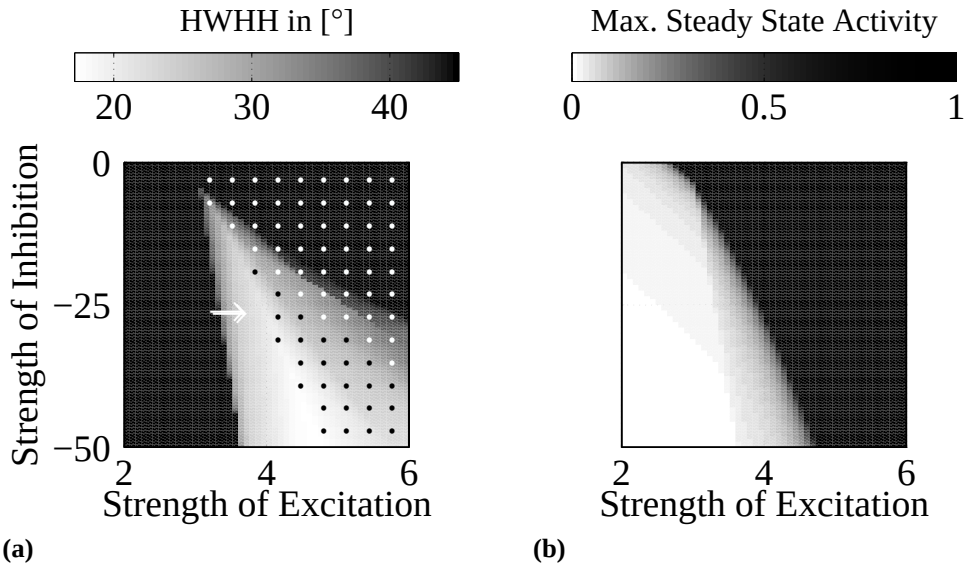


Figure 2.7: **(a)** Orientation tuning width (half width at half height) as a function of the strengths $W_C$ of excitatory and inhibitory connections. Gray values indicate tuning widths between 15° and 45°. The dots mark the area in parameter space where the responses of the cells saturate (the black region in b), at least at the optimal orientation. The white arrow marks the connection strengths which are typically used for the numerical simulations presented throughout this paper. **(b)** The maximal steady state cortical activity as a function of the strengths $W_C$ of excitatory and inhibitory connections. The gray values indicate the maximal steady state cortical activity. Stimuli were gratings with optimal spatial frequency and $c = 60\%$ contrast. The parameters were taken from Table 2.1 and Fig. 2.3.

if the postsynaptic membrane potential is close to the reversal potential (see Section 2.5). In both figures the tuning width and the cortical activity is coded by brightness values. In Figs. 2.7a and 2.8a dots indicate the region in parameter space for which the response of the cells saturates (equivalent to the black region where the maximal steady state activity is 1 in Figs. 2.7b and 2.8b).

The phase diagram shows three regimes for the model (cf. Sompolinsky and Shapley, 1997). If the recurrent excitation is not strong enough then the cortical response is not orientation selective and very weak. Towards stronger recurrent excitation, when the network amplifies the afferent input tuned response emerges (cf. Douglas et al., 1995). The tuning is sharpest if the depolarizing load (the geniculate and the lateral excitatory input) is just at the limit when it still can be balanced by the hyperpolarizing effects (lateral inhibition and the leakage) (cf. Tsodyks and Sejnowski, 1995). In other words, given a certain lateral inhibition strength, sharpest tuning emerges at the strongest lateral excitation when the cortical response still converges to a steady state. Note that because the number of excitatory connections is much larger than inhibitory ones (see Table 1), and the excitatory connections are less distributed among the orientation columns (they are more specific), balanced excitation and inhibition requires stronger inhibitory connections. Sharply tuned responses emerge when the effective strength of the single excitatory connections fulfill $W_C(e,\{e,i\}) \approx 3.1 + 0.03 \cdot W_C(i,\{e,i\})$,

a linear relationship. Thus the model predicts that sharp orientation tuning is robust against changes in the lateral activity as long as excitation and inhibition remain approximately balanced. If the visual cortex operates at high recurrent lateral excitatory and inhibitory load (cf., e.g. Peters et al., 1994) then the model predicts, that at a high, but not complete reduction of lateral activity, orientation tuning remains sharp. As a consequence, the cooling experiment by Ferster et al. (1996) is not necessarily decisive evidence for the afferent origin of the orientation tuning, assuming that inactivation of cortical activity was effective but not complete. Increasing the excitatory connection strength leads to saturation at optimal orientation, hence to broader tuning [4]. For reasons mentioned in the previous paragraph, close to the optimal tuning width the time to reach the steady state grows with the specificity of the response (data not shown). The phase diagram remains similar if iso-orientation inhibition is increased and cross-orientation inhibition is decreased (data not shown), but stronger excitation is needed to establish a tuned response and the region in parameter space for which cells do not saturate shrinks.

Figs. 2.8a shows the orientation tuning width as a function of different lateral connectivity patterns. The slopes of the excitatory and inhibitory connection percentages as a function of angular difference were taken as free parameters. The connectivity patterns are indicated in Fig. 2.8c for nine representative examples which cover the range of slopes shown in Figs. 2.8ab. Sharply selective response emerges in a wide regime where the slope of the percentage of connections as a function of angular difference for the excitatory connections is changed from $-6.6\,\frac{\%}{\mathrm{deg}}$ to $-0.2\,\frac{\%}{\mathrm{deg}}$ and the slope for the inhibitory connections is changed from $-0.2\,\frac{\%}{\mathrm{deg}}$ to $+0.4\frac{\%}{\mathrm{deg}}$. In a relatively large part of this regime the response saturates (dotted area) which again hints at the necessity to consider additional mechanisms to control runaway excitation for these lateral connectivity schemes. Sharpest tuning is achieved for fairly specific excitatory connectivity patterns and a slight dominance of iso- vs. cross-orientation inhibition. If the specificity of excitatory connections is decreased, stronger cross-orientation inhibition is required for a tuned response. If excitatory cells are connected in a highly specific pattern, then stronger iso-orientation inhibition is needed to control runaway excitation. If the excitation is too localized in the orientation domain then cell groups with close preferred orientation mutually suppress each other, and no tuned response emerges. Furthermore, extremely specific lateral inhibitory and excitatory couplings separate the different orientation columns and give rise to multimodal, periodic response patterns in the orientation domain (Carandini and Ringach, 1997) (data not shown).

### 2.3.3 Spatial frequency tuning and spatial receptive-fields

Fig. 2.9a shows plots of the half width at half height of the orientation tuning curves as a function of spatial frequency of the grating stimuli. Figs. 2.9b, c, d show the spatial frequency tuning curves for the afferent geniculate input to a single cell (thick line) and for the output at different stimulus orientations ($\Delta\alpha = 15°$, top to bottom). Orientation and spatial frequency tuning were investigated for four different lateral patterns of excitatory connections: highly specific (h) and strong (s) connections covering a distance of $500\mu m$, less specific (l) but strong (s) excitatory connections covering a distance of $500\mu m$, highly specific (h) but weak (w) connections covering a distance of $500\mu m$,

[4] Since cells have not been observed to saturate at their maximum firing frequency, mechanisms other than recurrent excitation and inhibition must be present to control runaway excitation if excitatory connections are too strong. The problem is connected to the problem of how to explain the saturation of the cortical neurons' contrast-response function (see, e.g. Albrecht and Hamilton, 1982). Candidate mechanisms include adaptation effects at the synaptic summation or spike generation, but they are still a matter of controversy.

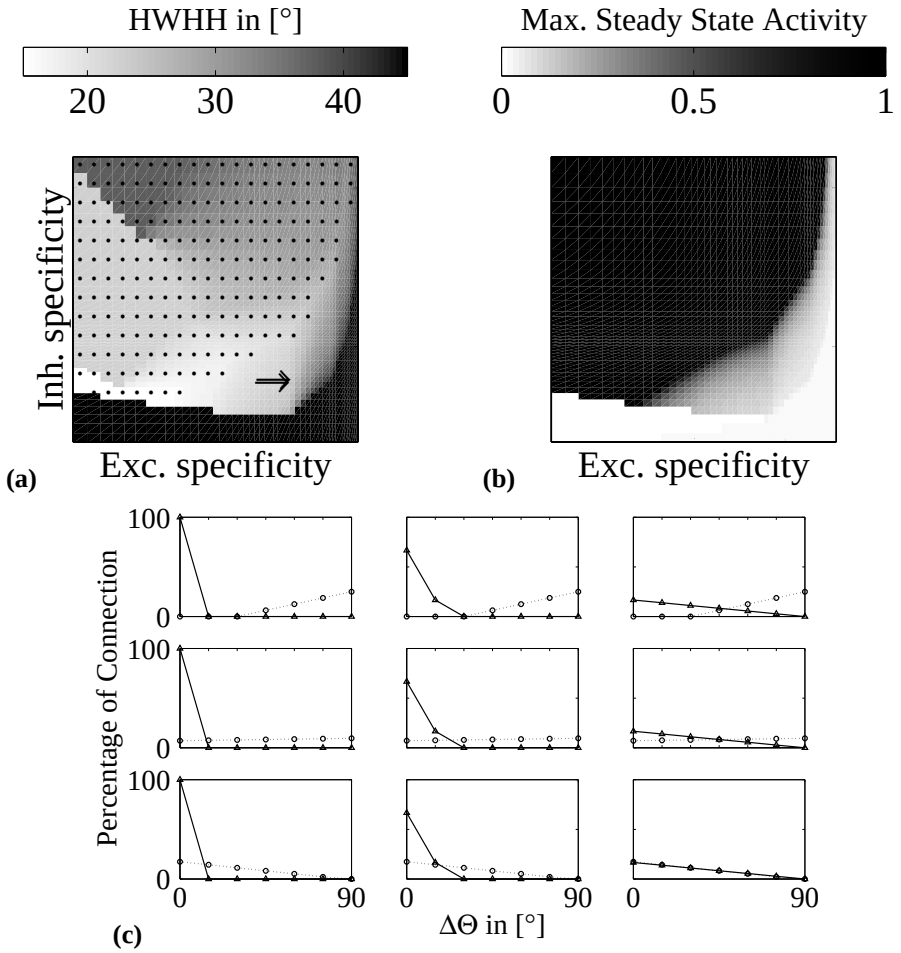


Figure 2.8: **(a)** Orientation tuning (half width at half height) as a function of the specificity of the lateral excitatory and inhibitory connections. The free parameters were the slopes $\frac{P(p,0)-P(p,15°)}{15°}$ of the percentages of excitatory and inhibitory lateral connections as a function of angular difference. Gray values indicate tuning widths between 15° and 45°, similar to Fig. 2.7. The arrow in (a) indicates the connectivity pattern which was typically used for the numerical simulations presented throughout this paper; dots mark the region in parameter space for which cells begin to saturate (the black region in b). **(b)** The maximal steady state cortical activity as a function of the strengths $W_C$ of the specificity of the lateral excitatory and inhibitory connections. The gray values indicate the maximal steady state cortical activity. Stimuli were gratings with optimal spatial frequency and $c = 60\%$ contrast. Parameters are given in Table 2.1, and Fig. 2.3. **(c)** Instead of labels in (a, b), to give an intuitive understanding of the varied parameters, we plotted nine representative connection percentage configurations of the excitatory (solid line) and inhibitory (dotted line) lateral connections as a function of angular difference. The plots correspond to the parameter values from the upper left corner via the midpoints to the lower left corner of the plots in (a, b).

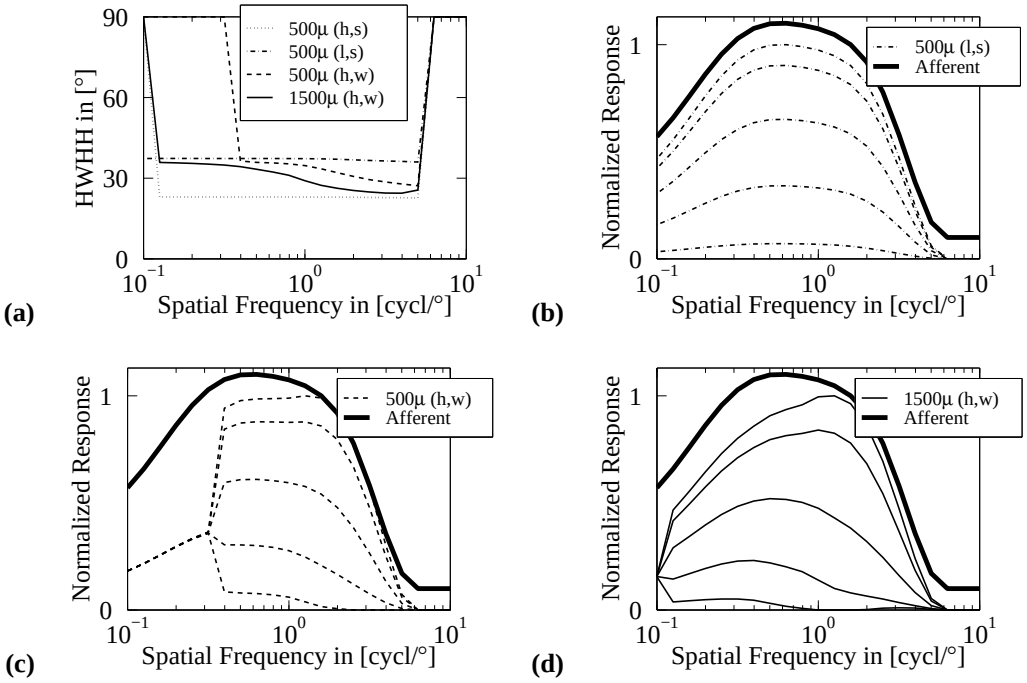


Figure 2.9: **(a)** The width of the orientation tuning curves (half width at half height) as a function of spatial frequency for four different cortical patterns of lateral connections. Parameters were (maximal connection distance; slopes of percentage of the excitatory and inhibitory connections as a function of the difference $\Delta\Theta$ in preferred orientation; the strength of excitation and inhibition $W_C$) 500 $\mu$m, -1.6 $\frac{\%}{\text{deg}}$, -0.06 $\frac{\%}{\text{deg}}$, 3.73, -25.4 (h, s); 500 $\mu$m, -0.06 $\frac{\%}{\text{deg}}$, -0.06 $\frac{\%}{\text{deg}}$, 4.33, -25.4 (l, s); 500$\mu$m, -1.6 $\frac{\%}{\text{deg}}$, -0.06 $\frac{\%}{\text{deg}}$, 3.33, -25.4 (h, w); 1500$\mu$m, -1.6 $\frac{\%}{\text{deg}}$, -0.06 $\frac{\%}{\text{deg}}$, 3.33, -25.4 (h, w). **(b,c,d)** Spatial frequency tuning curves for the afferent geniculate input to a single cell (thick line) and for the output at different stimulus orientations ($\Delta\alpha = 15°$, top to bottom). The figures (b,c,d) show results for the (500 $\mu$m, l, s), (500 $\mu$m, h, w), and (1500 $\mu$m, h, w) set of parameters, respectively. All spatial frequency curves are normalized to their maximum and the geniculate afferent input is shifted vertically up by 0.1 for clearer visualization. Stimulus contrast was $c = 60\%$; other parameters were taken from Table 2.1 and Fig. 2.3.

and highly specific (h) but weak (w) connections covering a distance of $1500\mu m$. Here "specificity" relates to the slope of the percentage of connections as a function of the difference $\Delta\Theta$ in preferred orientation; "strength" relates to the *total* amount of lateral excitation $W_C$ exerted by a single cell onto its targets.

While the preferred orientation is independent of spatial frequency, the width of orientation tuning may change (Fig. 2.9). As long as the total lateral excitation is strong enough the tuning width remains constant (Fig. 2.9a, dotted and dash-dotted line). For weaker lateral excitation (Fig. 2.9a, dashed and solid line), however, the width of the orientation tuning curve decreases with increasing spatial frequency as the initial columnar orientation bias grows. In this case the orientation tuning width and the cortical amplification depend on the initial orientation bias. This is in accordance with data obtained for cat (no measurements exist yet for layer 4C cells in macaque V1) (Webster and De Valois, 1985; Vidyasagar and Sigüenza, 1985; Hammond and Pomfrett, 1989). Note that this model prediction confutes previous claims by Troyer et al. (1998) that cortical amplifier models can not exhibit spatial frequency dependent orientation tuning. This effect is most pronounced for long lateral couplings (Fig. 2.9a, solid line) which integrate over a larger area in the visual field and for which a tuned response to lower frequencies can emerge even at weak excitation.

If the lateral excitation is sufficiently strong (Fig. 2.9b dash-dotted line) the peak frequency is independent of stimulus orientation and the tuning curve follows the spatial frequency tuning curve of the total afferent input $A$ (thick line). Although the specificity of the excitatory connections is low in this example, the increased value of the total excitation preserves tuning width for a wide range of spatial frequencies. In this region of parameter space the Fourier spectrum is polar-separable, i.e. it can be decomposed into an orientation tuning curve multiplied by a spatial frequency tuning curve. While polar-separable receptive-field profiles in the frequency domain have been reported for a large group of visual cortical simple cells in cat area 17 (Webster and De Valois, 1985), no detailed experiments have yet been done in macaque V1.

For the case of weak lateral excitation (Figs. 2.9c,d) the shape of the spatial frequency tuning curves depends on stimulus orientation. The peak spatial frequency decreases if stimulus orientation differs from the optimal value leading to non polar-separable tuning curves. Note that in contrast to models using elongated geniculo-cortical projection patterns, the spatial frequency tuning curves of the cortical cells affected by the lateral input, and differ from the spatial frequency tuning of their total geniculate input. Fig. 2.9c also shows that below a critical frequency, whose value depends on the spatial extent of the lateral interactions, the initial orientation bias is too small to be amplified and no orientation tuning emerges. As a consequence of this low frequency cutoff, spatial frequency tuning of cortical model cells is sharper than the tuning of the afferent geniculate cells (cf. De Valois et al., 1982; Hawken and Parker, 1987). The larger the spatial extent of the lateral interactions, the lower is the critical value for the spatial frequency, and the spatial frequency tuning of cortical cells eventually approaches the tuning properties of the LGN (Fig. 2.9d).

In summary, the *strength* of the recurrent excitation determines if the orientation tuning depends on the initial orientation bias, hence it determines if the two-dimensional spatial frequency spectra are polar-separable. The *specificity* of the recurrent connections determines the orientation tuning width and the *range* of the anisotropic recurrent excitation affects the size of the initial columnar orientation bias and thus the robustness of the orientation selectivity.

In the following, we describe the receptive-field profiles of the modeled cortical cells by two different experimental paradigms (cf. Jones et al., 1987; Jones and Palmer, 1987*a*; Jones and Palmer, 1987*b*). Simulating the first paradigm, the receptive-field profile was calculated (Figs. 2.10a,d) by

presenting stationary sinusoidal gratings[5] with various values of horizontal ($k_1$) and vertical ($k_2$) spatial frequency components, and the total input $I$ (Eq. 2.6) arriving to a modeled cortical cell was calculated. The corresponding receptive-field profile pattern in spatial domain (Figs. 2.10b,e) was then computed by inverse Fourier transformation. In the second paradigm, the total cortical input $I$ was calculated for small, circular symmetric Gaussian spots which were presented at different locations in the visual field (Figs. 2.10c,f).

The resulting receptive-field profiles resemble simple cells and are composed of an elongated excitatory center flanked by two elongated suppressive domains. receptive-field size in the grating-paradigm decreases for higher contrast but the receptive-fields remain larger than when measured in the spot-paradigm (cf. Jones and Palmer (1987*a*) for cat area 17)[6]. This effect is a consequence of the finite thresholds of the thalamic and cortical transfer functions. At low contrast geniculate and cortical cells do not respond to high spatial frequencies and the total input of the cells remains below threshold. Consequently, receptive-fields are small in the frequency domain and become large when transformed back to the spatial domain. At higher stimulus contrast the cells operate mainly in the linear regime of their transfer function and respond also to higher frequencies. Therefore, as the response profile in the frequency domain increases, the resulting receptive-field in the spatial domain shrinks (cf. Schumer and Movshon, 1984; Jagadeesh and Ferster, 1990), and the receptive field profile becomes closer to the profiles determined by spot stimuli. The aspect ratio of the receptive-fields (length / width) is similar for both experimental paradigms and it is independent of stimulus contrast, but the receptive-field profiles determined by spot stimuli predict a broader width of the orientation tuning curve at low spatial frequencies (Volgushev et al., 1996) because they are smaller than the receptive-fields obtained by the grating-paradigm. The diameter of the receptive-field determined by the spot paradigm is 0.2-0.4°, and it fits to the experimental data (Hubel and Wiesel, 1974; Blasdel and Fitzpatrick, 1984). The differences in the receptive-field profiles and orientation selectivity obtained by the different measurement methods indicate that simple cells cannot be fully described by linear spatial filters. Rectification by the transfer function changes the response to gratings of different spatial frequencies and thus the receptive-field structure in the space domain.

The inhibitory subfields in the receptive-field profiles are generated by lateral inhibition, and the model predicts that blocking inhibition erases the subfield structure (Sillito, 1975). Similarly, the elongation of the receptive fields is generated by the anisotropic lateral excitatory connections, thus the model predicts that by blocking lateral excitation, the receptive-field profiles become more and more circular. The model also predicts that the local distortions in the visuotopic map investigated systematically by Das and Gilbert (1997) could be due to the strong lateral recurrency. The thalamocortically defined receptive-field shape can be changed and shifted by the recurrent connections. Strong recurrent excitatory connections result in a shift of the receptive-fields towards the "center of mass" of the strongly connected group of activated cells. Hence small clusters of cells with similar orientation preference emerge, which are characterized by strongly overlapping receptive-fields.

Additionally to test whether our simplifying assumption of neglecting ON-OFF interactions

[5] As the simulated experiment assumed linearity, we took into account only the significant response to a cosine grating with zero phase with respect to the cell's receptive-field center and neglected the small response to sine grating which is only due to the nonlinear rectification of the modeled cells, and in the case of an idealistic linear cell should be zero. This approximates the original experiment and implicitly takes into account the model assumption that the cortical cells targeting the investigated modeled cell occupy a point symmetric region in the visual field with respect to the cell's receptive-field center. Hence, the receptive-field profile of that cell is expected to be point symmetric as well.

[6] No measurements exist yet for the two-dimensional spatial profile of layer 4C cells in macaque V1.

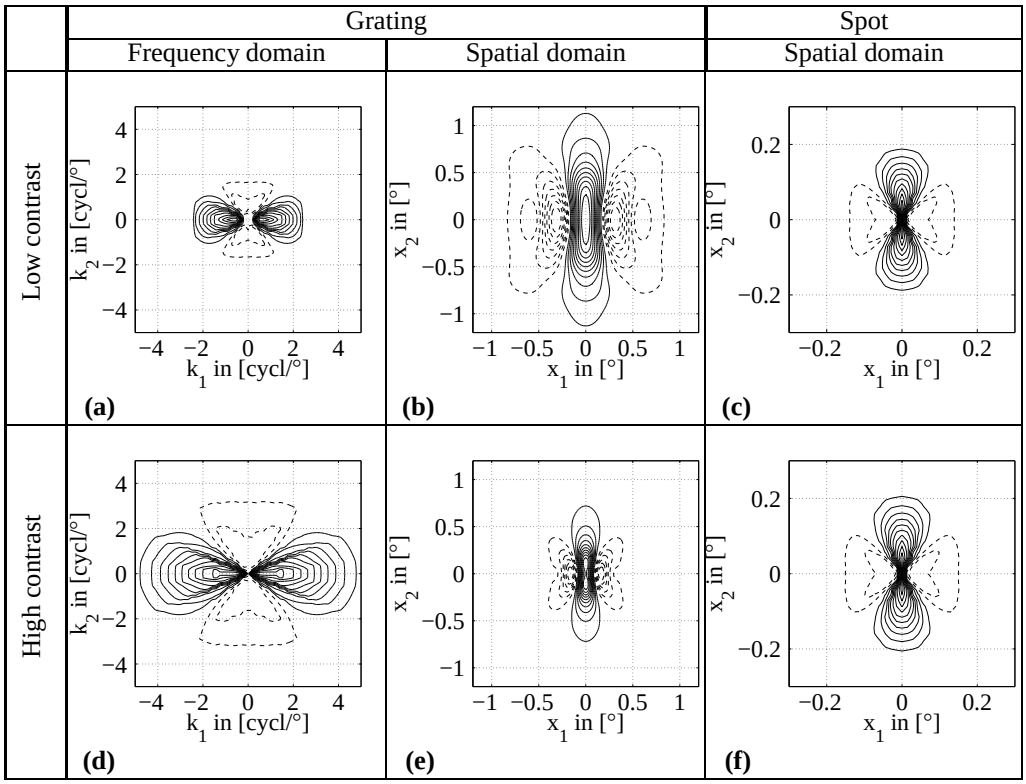


Figure 2.10: Contour plots of the spatial receptive-field for an excitatory cell which is tuned to $\Theta = 90°$. The contour lines mark intervals of 10% of the maximal positive input value. Solid and dashed contours denote excitatory and inhibitory subfields, respectively. **(a,d)** receptive-field profiles in frequency domain was measured via grating stimuli of different spatial frequencies and orientations. Contours indicate lines of constant total input for cosine gratings with different spatial frequencies along the horizontal ($k_1$) and the vertical ($k_2$) axis. Stimulus (Michelson) contrast was $c = 4\%$ in (a) and $c = 60\%$ in (d). **(b,e)** receptive-field profiles in spatial domain. The patterns were computed by an inverse Fourier transform of the patterns shown in (a,d). **(c,f)** Receptive field profiles in spatial domain calculated by small spot stimuli with $c = 60\%$ (c) and $c = 100\%$ (f) Weber contrast. Spot size was $s_1 = 0.04°$, $s_2 = 0.04°$ (Eq. 1). The stimulus contrast was chosen such that the maximum total input for a bar contrast of 60% (100% ) roughly corresponds to the maximum total input for a grating contrast of 4% (60% ). Note the different scale factors in (b,e) and (c,f).

modifies qualitatively the behavior of the modeled simple cells we calculated the receptive-field profiles including OFF pathways too (data not shown). In the simulations the OFF geniculate cells project to a cortical simple cell from a circularly symmetric ring which overlaps slightly with the inner ON center. The simulations lead to receptive-fields very similar to those calculated only with the ON pathway. The included OFF projections do not alter the cells' nonlinear behavior, because it

is a simple consequence of the firing threshold.

## 2.4 Discussion

We have investigated a model for the generation of orientation preference and tuning and demonstrated that it is not essential to have orientation biased afferent inputs to individual cortical cells; rather cortical cells can acquire an orientation bias via anisotropic excitatory lateral intracortical connections. Our model predicts that if the orientation bias is both generated and amplified via recurrent excitation and inhibition, the emergent response properties are quite similar to the response properties which are observed experimentally. Because tuning properties are determined to a large extent by recurrent amplification it is not surprising that model predictions are quite similar also to the predictions of cortical amplifier models based on a Hubel and Wiesel style afferent orientation bias. Therefore, the hypothesis of an intracortical orientation bias due to lateral excitatory connections is a sensible alternative to the notion of an afferent bias due to elongated geniculo-cortical projection patterns, and experiments must yet be designed to decide between both hypotheses. In the following we will discuss model assumptions and model predictions in more detail.

### 2.4.1 Model assumptions

We tried to involve as few mechanisms as seemed to be necessary to explain orientation selectivity, which allowed us to have an understanding of the behavior of a "clear-cut" system instead of having mixed effects originating from a mixture of assumptions. (i) For this we did not include an additional afferent orientation bias, which—as we discuss later (see Section 4.2)—is very likely to be present. This restriction allowed to demonstrate that anisotropic lateral excitatory connectivity itself is sufficient to generate the initial orientation bias and sharp tuning. (ii) To avoid assumptions regarding the arrangement of ON and OFF center geniculate cells in the projective field to a simple cell in visual cortex the presented model framework contains only ON geniculate projections. However, additional simulations showed (see Section 3.3) that OFF projections arranged in a circular symmetric ring around the ON geniculate center—in accordance with biological findings (see Section 2.1)—do not alter the emergence of sharply orientation tuned response, but rather increase robustness of tuning. (iii) The cortical circuit is modeled within a mean-field-framework. The drawback of this approach is that it is hard to relate the state variables directly to physiological properties of real neurons. To avoid over-interpretation of the neural activity dynamics, neural activity and time is given in arbitrary units in our presentation. The mean-field approach was motivated by computational efficacy and it gave the possibility of direct comparison of our phase-diagrams to previous results (Ben-Yishai et al., 1995; Sompolinsky and Shapley, 1997). The simplified simple cell model also demonstrated clearly the necessity for additional mechanisms for contrast saturation and gain control. Exploration of dynamics of orientation tuning and contrast processing within a spiking framework have been done in our accompanying studies (Adorján et al., 1998; Stetter et al., 1998).

### 2.4.2 Intracortical vs. afferent origin of the orientation bias

In a recent study of orientation selective cells in macaque V1 Ringach et al. (1997*a*) report orientation tuning dynamics which leads from an initially orientation nonspecific to a sharply tuned response within 10-15 msec for some cells in layer 4C$\alpha$. This study provides a strong indication for

our hypothesis that intracortical factors play an important role in generating orientation selectivity in macaque monkey V1. While the large stimuli they used may have recruited influences from beyond the core orientation-selective mechanism of their cells, this study nonetheless provides strong evidence for our hypothesis that intracortical factors play an important role in generating orientation selectivity in macaque V1. We note, however, that other factors, such as delays in propagation of different synaptic potentials which might manifest themselves as different response timings across the receptive-field (Reid et al., 1991; Livingstone, 1998), could be responsible for the temporal dynamics of orientation tuning.

Although LGN receptive-fields are rarely perfectly nonoriented (Smith et al., 1990), given the need for retinotopy and the substantial overlap amongst individual geniculo-cortical terminal fields, it is hard to imagine how these slight orientation biases of single LGN cells could be reflected in the aggregate postsynaptic input to cortical neurons. While such a mechanism might well serve to generate orientation selectivity in other species, it seems less likely in the primate. It seems odd to us that one portion of the thalamic input zone (layer 4C$\beta$) should lack orientation selectivity, while another (4C$\alpha$) has it in abundance. This difference is accompanied by a difference in lateral connectivity supporting our suggestion that it is the lateral intrinsic cortical connections that might generate orientation selectivity in the primate.

A much larger body of physiological data are available for cats and ferrets but experimental data are partially contradictory. Pei et al. (1994) and Volgushev et al. (1995) report orientation tuning dynamics with an initially unspecific "bump" of excitatory input followed by an amplification of the optimally oriented cells' output and a shutdown of the activity of cells tuned to the null orientation. Other groups, however, report highly orientation specific EPSPs for cats (Ferster, 1986; Douglas et al., 1991). The experimental results, however, should be interpreted with some care. Douglas et al. (1991) showed sharply direction specific excitation on cortical cells in cat primary visual cortex, although, as they mention, the cell received non direction specific input from the LGN. This indicates that it may not has been possible to fully separate the small afferent excitation from the intracortical contributions. Furthermore, small nonspecific EPSPs present for a few milliseconds do not necessarily evoke spikes (Adorján et al., 1998), hence it could be difficult to detect an initially unspecific response via single cell spike trains (Celebrini et al., 1993). Cross-correlation techniques (applied for monkeys) (Ringach et al., 1997*a*; Ringach et al., 1997*b*) on the other hand should be more reliable because they accumulate nonspecific neural responses and thus show more realistic orientation tuning dynamics. Strongest evidence for an afferent orientation bias in cat is provided by the study of (Ferster et al., 1996), although Vidyasagar et al. (1996) questioned whether all local intracortical circuits and all cortical afferents to area 17 were truly silenced by the cooling procedure.

In summary, the experimental facts remain somewhat ambiguous for cat and are very sparse for primates, hence our model study should be regarded as an alternative hypothesis to the idea of an afferent orientation bias which now needs to be explored experimentally. Though one major outcome of our modeling study is that the predictions of the "intracortical hypothesis" are quite similar to the predictions of the previous cortical amplifier models there are important differences. $(i)$ The tuned response gradually emerges after stimulus onset with a nonspecific small initial excitatory response which depends on the strength of lateral excitation (Fig. 2.6). $(ii)$ The orientation tuning completely vanishes when the recurrent excitation is perfectly blocked. $(iii)$ The receptive-field profiles in the spatial frequency domain are polar-separable and the spatial frequency tuning of the cortical cells follows the geniculate input to cortical cells if lateral excitation is strong. The response profiles in the spatial frequency domain are non polar-separable and the cortical spatial frequency tuning is different from the spatial frequency tuning of the geniculate input to cortical cells at weak lateral

excitation (Fig. 2.9).

It could well be that different species have implemented orientation selectivity in different ways; after all, the rabbit generates orientation specificity within its retina (Bloomfield, 1994). It could also be that elongated convergence of geniculate arbors and anisotropic lateral connections parallel to the preferred orientation are *both* present in different species, and jointly contribute to the emergence of an orientation bias. This would explain the existence of orientationally biased EPSPs and at the same time explain the results of the developmental experiments (Kim and Bonhoeffer, 1994; Bonhoeffer and Goedecke, 1996) via the existence of a cortical "trace". The addition of an afferent bias as long as the orientation preferences are similar would serve as a cooperative mechanism. If the recurrent lateral excitation is strong, and thus the orientation tuning width is independent from the initial orientation bias, it would increase the robustness of the tuning to noise. If the lateral excitation is not strong enough to ensure bias independent tuning width, an additional afferent orientation bias would sharpen the tuning. The addition of an afferent bias which is different from the lateral bias would weaken the orientation bias a cortical cell receives; depending on the strength of the lateral excitatory feedback this would decrease the noise robustness and/or increase the tuning width. However, this case is unlikely to happen because during development of the cortical maps, cells with similar tuning properties tend to build strong connections.

### 2.4.3 Sidestep connections and orientation bias

Layer 4C in macaque striate cortex contains spiny stellate cells with axon collaterals which project laterally along one axis and which terminate in one or several patches (Lund, 1987; Anderson et al., 1993; Yoshioka et al., 1994). Cells of this kind are absent in lower 4C$\beta$ where neurons lack orientation selectivity, but are found in mid 4C and upper 4C$\alpha$ where orientation selectivity emerges (Blasdel and Fitzpatrick, 1984; Hawken and Parker, 1984). The number of cells with sidesteps increases from the middle to the upper part of the layer and correlates with increasing distance of the lateral spread of the sidesteps, and with increasing number of orientation selective cells and with sharpness of orientation tuning. If these cells are indeed responsible for the orientation bias, lateral sidesteps have to link neurons with co-oriented and co-axially aligned receptive-fields. Studies on stepped connections (Bolz and Gilbert, 1989; Schwartz and Bolz, 1991; Bosking et al., 1997) support this idea in general but the data relate to different species (cat, tree-shrew) and to different layers (5, 6 and 2/3). No data of this kind exist yet for layer 4C in macaque V1, hence the above assumptions should be regarded as model predictions. Note, however, that in the model the width of the orientation tuning curve depends on the range of the lateral sidesteps for weaker recurrent excitation (Fig. 2.9) which could explain the observation that increasing distance of the lateral sidesteps in upper layer 4C$\alpha$ correlates with increase in sharpness of orientation tuning in depth of the layer.

Are lateral sidesteps, similar to the axon collaterals actually found in layer 4C, necessary to explain the generation of an intracortical orientation bias? "Sidestep" axon collaterals in upper layer 4C$\alpha$ project laterally over a distance of 1500 $\mu$m from the injected point i.e. 3 mm total diameter which corresponds to a distance of $1.2^o$ in the visual field at $5^o$ degrees eccentricity. This corresponds to four times the diameter of a receptive field of a non-oriented unit. This offset leads to a fairly large and robust bias. Depending on the magnitude of neuronal noise, however, a less strong bias may also suffice, for example a bias generated by a lateral offset of receptive fields which is of the order of the receptive-field scatter in layer 4C (300-400 $\mu$m, (Hubel and Wiesel, 1974; Blasdel and Fitzpatrick, 1984)). If this is the case the anatomical correlate of the orientation specific lateral excitatory connections would be harder to detect.

For the lateral connectivity our model predicts that those cells are connected together by strong lateral excitatory connections whose receptive fields as determined by the thalamic input alone are shifted along the axis of their preferred orientation. In other words, (i) the projection patterns of cortical excitatory cells mapped back to the visual field coordinate system are elongated and (ii) cells with parallel elongation in their projection patterns are more likely to be connected by excitatory connections. Note, that the shape of the patches of the strongly connected cells or the lateral connectivity pattern is not necessarily similar to the shape of the area they cover in the visual space, because of the slight scatter in the cortical retinotopy (Hubel and Wiesel, 1974; Blasdel and Fitzpatrick, 1984; Das and Gilbert, 1997). The functional consequences of the anisotropy of the projection pattern are: (i) orientation bias is generated intracortically so that the preferred orientation of the cells is parallel to the axis of elongation and (ii) the lateral excitation is highly orientation specific.

One approach to test whether sidestep projections could subserve generation of orientation selectivity would be to use small injections of the retrograde tracer cholera toxin, injecting small points in layer 4C, and examining the pattern of retrogradely labeled cells both in the LGN and within cortex. This could indicate the aspect ratio of connectional fields providing input to a point in 4C. Furthermore, LGN afferents terminate in both the alpha and beta divisions of layer 4C, while orientation selectivity is essentially restricted to the middle and upper portions of the layer. Thus, if our model is correct, one might expect to find the kinds of local circuits we describe to be missing from the lower portion of layer 4C. Electrophysiological studies might also be used to test our model. We have earlier cited pharmacological studies that inactivated inhibitory circuits within cortex. If suitable methods could be found to selectively inactivate *cortical excitatory* circuits, while leaving excitatory geniculo-cortical transmission unaffected, one could distinguish between the afferent and intracortical hypotheses. We have earlier suggested that blocking local cortical excitation might also make local receptive-fields more circular in shape, and could eliminate the local distortions in the retinotopic map of visual space. One could also measure whether cells in 4C are capable of responding to multiple orientations at a single retinotopic position (Carandini and Ringach, 1997). Such studies could also provide evidence to distinguish between purely afferent and recurrent intracortical hypotheses.

# Chapter 3

# Generating orientation selectivity intracortically—a statistical neural field approach


## Abstract

We apply the recently proposed statistical neural field approach (Grőbler et al., 1998; Barna et al., 1998) for modeling orientation selectivity in the primary visual cortex. Firstly, we demonstrate that the neural field approach is a powerful tool for modeling neural structures with specific lateral connections. Secondly, we test in a biologically more plausible way our hypothesis (Adorján, Levitt, Lund and Obermayer, 1999) that orientation bias and tuning in macaque striate cortex can be generated by the same lateral interactions (see previous chapter and Adorján, Levitt, Lund and Obermayer, 1999). The spiking neural model shows that (i) contrast invariant tuning emerges, and (ii) the tuning dynamics of the membrane potential and the firing rate are in accordance with observations which until now seemed to be contradictory (Celebrini et al., 1993; Pei et al., 1994; Ringach et al., 1997*a*; Pernberg et al., 1998).[1]


## 3.1 Introduction

The emergence of orientation selectivity in the primary visual cortex of higher mammals has been an active area of research during the past decades. In contrast to previous models, where the initial orientation bias is generated by convergent geniculate input to simple cells and subsequently sharpened by lateral circuits, we showed that the initial orientation bias can be *both* generated and amplified intracortically via the same anisotropic recurrent excitatory and inhibitory connections (previous chapter and Adorján, Levitt, Lund and Obermayer, 1999). This is referred to as the "intracortical hypothesis".

The "intracortical hypothesis" addresses the geometrical relation between the recurrent cortical

[1] This chapter is based on (Adorján et al., 1998).

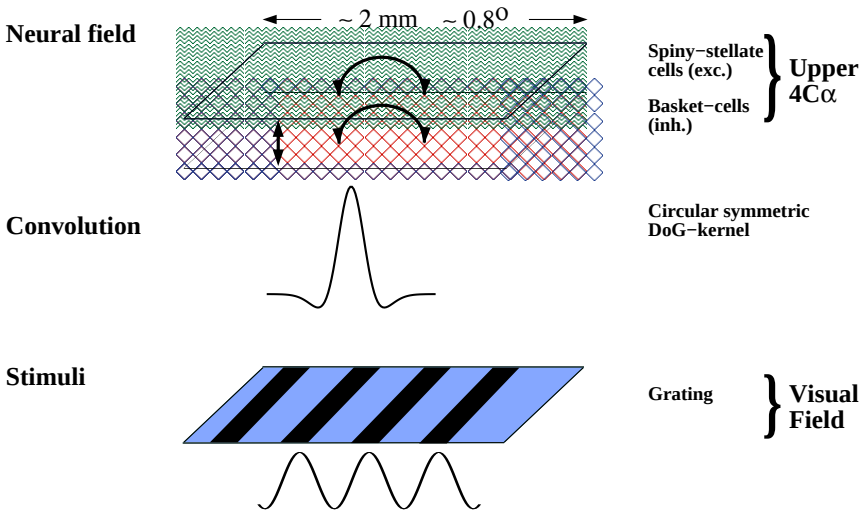


Figure 3.1: The structure of the connectionist model for generating orientation selectivity intracortically.

connectivity and the receptive field positions of the simple cells in V1. It suggests that the recurrent excitatory coupling is strongest between cortical cells whose receptive fields are located on a particular axis in the visual field (cf. Fig. 2.3) such that the sum of the total geniculate input arriving to all such coupled cells depends on the orientation of the bar or grating stimulus (*columnar orientation bias*). This anisotropy generates the initial orientation bias that is further amplified and sharpened by the same lateral circuitry. Here we re-investigate the "intracortical hypothesis" in a framework that allows the modeling of large scale neural networks with realistic geometrical arrangement of the recurrent connections. We study how anisotropic the recurrent connectivity should be in order to generate an initial orientation bias that is strong enough to overcome the inherent noise in the network.

Furthermore, the present investigation focuses also on the dynamics of the membrane potential and the spiking activity as a function of orientation. Our previous study highlighted that the initial orientation unspecific activation after stimulus onset could be an indication of the strong intracortical involvement in the generation of orientation selectivity (cf. section 2.4.2 and Fig. 2.6). The present statistical neural field model is based on a more realistic leaky integrator single cell model that allows the investigation of the membrane potential and the spiking dynamics. We attempt to provide an explanation for the experimental data that show immediate and sharp tuning of single cell spiking activity (Celebrini et al., 1993; Pernberg et al., 1998), but shows a gradual sharpening of orientation tuning of the membrane potential (Pei et al., 1994; Pernberg et al., 1998), and the spiking dynamics that is measured for several trials (Ringach et al., 1997*a*).

## 3.2 The statistical model

The three layer computational model (Fig. 3.1) represents the visual field, the spatial filtering by the lateral geniculate (LGN) cells, and a layer of orientation selective simple cells in the thalamic recipient zone in upper layer 4 in primate visual cortex. Stimuli are stationary sinusoidal gratings with a diameter of 0.6°. The feed-forward input to the cortical layer is modeled by the rectified convolution of the stationary sinusoidal stimulus with the circular symmetric receptive field profiles parameterized according to (Spear et al., 1994). For the feed-forward input to the cortical layer we use the same model with the same parameters as in the previous chapter and in (Adorján et al., 1998).

The cortical layer in the current neural field model consists of excitatory and fast inhibitory neuron populations with number ratio of 4:1, and a density of 10000 neurons per $mm^2$. Each neuron has 5000 synaptic connections from which 10% originate from feed-forward projections. Following (Ventriglia, 1974; Grőbler et al., 1998; Barna et al., 1998) we describe the population activity of neurons as well as the propagation of spikes in terms of probability density functions (p.d.f.). (A similar statistical framework for modeling orientation selectivity has been recently developed independently (Nykamp and Tranchina, 2000).) For a neuron of type $s$ at each spatial location $r$ in the neural field the probability density $g_s(r,u,t)$ of being at a sub-threshold membrane potential $u$ at time $t$ is described via a single compartmental model. When a cell fires, it emits as many "spikes" as it has synaptic terminals. The probability density $f_s^{\alpha}(r,t)$ of spikes emitted by a neuron of type $s$ being at a position $r$ in the neural field at time $t$ traveling at direction $\alpha$ is then determined as following. The spikes propagate radially in all directions and diffuse in space simulating the effect of different propagation delays of action potentials. Synaptic connections are modeled by absorption of spikes at each point in the neural field, changing the membrane potential, hence the probability density of the neurons at this point. A novel feature of the current model framework is that specific recurrent connections are implemented by anisotropic spike emission and absorption distributions (see section 3.2.3).

In the following sections we describe the neural field model for the cortical layer in detail. The model parameters are endetailed in table 3.1. The model parameters correspond to a membrane time constant of 14 msec for excitatory and 7.7 msec for inhibitory neurons. In the model the EPSP amplitude is $0.1-0.4$ mV, and the IPSP amplitude is 0.5 mV at resting potential. The model with the enlisted passive physiological parameters approximate well the behavior of regular-spiking (excitatory) and fast-spiking (inhibitory) neurons (McCormick et al., 1985; Komatsu et al., 1988; Stratford et al., 1996).

### 3.2.1 Neurons

The time evolution of the probability density $g_s(r,u,t)$ of being at a sub-threshold membrane potential $u$ at time $t$ is

$$\frac{\partial}{\partial t}g_s(\mathbf{r},u,t)+\frac{\partial}{\partial u}\Big(\overbrace{\varepsilon_s(\mathbf{r},u,t)}^{\text{leaky integrator}}\ g_s(\mathbf{r},u,t)\Big)-\overbrace{\frac{D_u}{2}\cdot\frac{\partial^2}{\partial u^2}g_s(\mathbf{r},u,t)}^{\text{diffusion}}= \\ =\underbrace{b_s(\mathbf{r},u,t)}_{\text{rate of returning from refr.}}-\underbrace{n_s(\mathbf{r},u,t)}_{\text{rate of firing}}. \tag{3.1}$$

| Property | Exc. | Inh. |
|---|---|---|
| Resting potential $E_{\text{rest}}$ [Vm] | $-65$ | $-65$ |
| Membrane capacitance $C$ [nF] | 0.5 | 0.2 |
| Leakage conductance $\gamma_{\text{leak}}$ [nS] | 35 | 26 |
| Firing threshold $\Theta$ [mV] | $-40$ | $-30$ |
| Absolute refractory period $T^{\text{refr}}$ [msec] | 3 | 2 |
| Membrane potential diffusion constant $D_u$ [mV$^2$/msec] | 4 | 4 |
| Postsynaptic resting potential $E_{s's}$ [mV] | $-5$ | $-80$ |
| Maximal excitatory postsynaptic conductance $\overline{\gamma}_{es}$ [nS] | 0.55 | 0.8 |
| Maximal inhibitory postsynaptic conductance $\overline{\gamma}_{is}$ [nS] | 7.5 | 4 |
| Time to peak $\tau_{\text{peak}}$ [msec] | 1 | 1 |
| Velocity of spike propagation $v$ [mm/msec] | 0.5 | 0.5 |
| Radial diffusion constant $D_r$ [mm$^2$/msec] | 0.034 | 0.034 |
| Number of cells | 32000 | 8000 |
| Number of excitatory synapses | 4000 | 1000 |
| Number of inhibitory synapses | 4000 | 1000 |

Table 3.1: Summary of the model parameters of the statistical neural field model.

The left-hand side describes the inter-spike dynamics that balances with a source and a sink term. The sink term corresponds to the rate of spike emission (firing) $n(r,u,t)$

$$n_s(\mathbf{r},u,t) = \begin{cases} \varepsilon_s(\mathbf{r},u,t)\cdot g_s(\mathbf{r},u,t) & \text{if } \varepsilon_s(\mathbf{r},u,t) > 0 \text{ and } u \geq \Theta_s \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

where $\Theta_s$ is the firing threshold. The source term corresponds to the rate of returning from the refractory period is

$$b_s(\mathbf{r},u,t) \qquad = \int_{-\infty}^{t} dt' \int_{-\infty}^{\infty} du'\, n_s(\mathbf{r},u',t') \cdot \delta(u - U_s^{\text{ret}}) \cdot \delta(t - (t' + T^{\text{refr}}))\,, \tag{3.3}$$

where $T^{\text{refr}}$ is duration of the absolute refractory period and $U_s^{\text{ret}}$ is the reset potential.

Analogous to a leaky integrator model, the rate of change $\varepsilon_s(\mathbf{r},u,t)$ in the sub-threshold membrane potential $u$ is

$$C_s \varepsilon_s(\mathbf{r},u,t) = -\sum_{s'} \gamma_{s's}(\mathbf{r},t)\cdot(u - E_{s's}) - \gamma_{\text{leak}}\cdot(u - E_{\text{rest}})\,. \tag{3.4}$$

The postsynaptic conductance $\gamma_{s's}(\mathbf{r},t)$ is calculated by the convolution of the presynaptic events with the alpha-function Eq. 3.7. Noise is taken into account by the additive diffusion term. The rate of firing $n(r,u,t)$ is proportional to the probability of firing at membrane potential u and the rate of returning from the refractory period is given by $b(r,u,t)$.

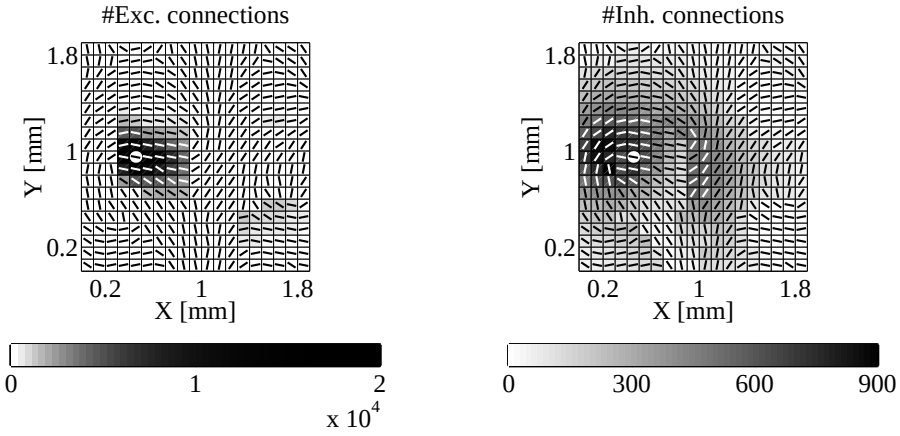


Figure 3.2: The excitatory and inhibitory recurrent connectivity. The anisotropy of the recurrent connections from each spatial element is denoted by the angle of the overlaid bars. The bars are inverted only for visualization reasons. Gray values indicate the number of recurrent connections originating from the spatial element marked by a white dot. The image is obtained by calculating the propagation of a spike population (Eq. 3.5) emitted by the element marked by a white dot with the anisotropic emission $\lambda(|\alpha - EL(r)|)$ and the absorption $\sigma(|\alpha - EL(r)|)$ distributions. For this simulation the neural activity was set to zero to obtain the connectivity pattern only for one spatial element.

### 3.2.2 Spikes

The dynamics of the probability density $f_s^{\alpha}(r,t)$ for the spikes at the direction $\alpha$, position $r$ and time $t$ is

$$\frac{\partial f_{s'}^{(\alpha)}(\mathbf{r},t)}{\partial t} + \overbrace{(\mathbf{v}_{s'}\nabla) f_{s'}^{(\alpha)}(\mathbf{r},t)}^{\text{drift}} - \overbrace{\frac{D_r}{2}\cdot\nabla^2 f_{s'}^{(\alpha)}(\mathbf{r},t)}^{\text{diffusion}} = \\ -\underbrace{\sigma_{s'}^{(|\alpha-EL(r)|)} f_{s'}^{(\alpha)}(\mathbf{r},t)}_{\text{absorption}} + \underbrace{\lambda_{s'}^{(|\alpha-EL(r)|)}\cdot\int_{-\infty}^{\infty} du' n_{s'}(\mathbf{r},u',t)}_{\text{emission}} . \tag{3.5}$$

The change in probability density, a drift with velocity $v(\alpha)$ and a diffusion in space which accounts for different propagation delays balances the rate of spike absorption and spike emission. The velocity of the spikes emitted by population $s$ is $v_s$; the radial diffusion constant is $D_r$. The anisotropic recurrent connections are determined according to the absorption $\sigma_{s'}^{(|\alpha-EL(r)|)}$ and the emission $\lambda_{s'}^{(|\alpha-EL(r)|)}$ distribution as described in the next section.

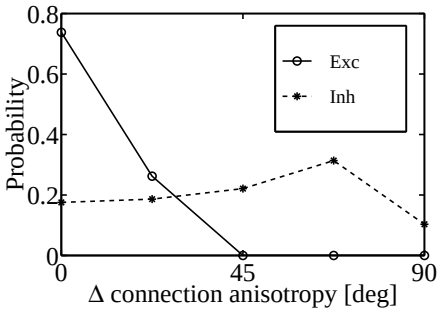


Figure 3.3: Probability of recurrent excitatory (solid line) and inhibitory (dashed line) connections *vs.* difference in projection anisotropy. The excitation is highly specific with respect to the projection field anisotropy $EL(r)$. In contrast, inhibitory connections are not specific.

### 3.2.3 Recurrent connectivity

The main assumption and hypothesis of the current work is that orientation selectivity is generated purely intracortically by the anisotropic recurrent connections. Here we describe how the specific anisotropic connectivity pattern is implemented in the neural field framework.

A label $EL(r) = [0 \dots 180]$ that determines the orientation of the anisotropy of the neurons' projection pattern is assigned to each discrete element at different positions $r$ (see the bars in Fig. 3.2). We assume that *(i)* the projection patterns mapped back to the visual field coordinate system of the cells are elongated and *(ii)* cells with parallel elongation in their projection patterns are more likely to be connected by excitatory connections. The resulting connection maps for discrete spatial element is showed in Fig. 3.2. Note again, that the connection map is defined and plotted in the visual field coordinate system that is equivalent to the cortical coordinate system if and only if the retinotopy is perfect.

We assume less anisotropic lateral inhibitory connections, and stronger inhibitory interactions between cells whose projection fields are orthogonal to each other. Because the generated orientation map is smooth, and the number of connections decreases exponentially with distance from the cell body we effectively obtain approximately equally distributed, non orientation specific recurrent inhibition. The connection specificity as a function of elongation of the projection pattern $EL(r)$ is shown in Fig. 3.3. The excitatory connections are highly specific, while the inhibitory connections are unspecific. By simulations we will demonstrate (section 3.3) that the two assumptions about the connection anisotropy have the following consequences: the preferred orientation of the cells will be parallel to the orientation of the elongation of their projection patterns, and lateral excitation will be orientation specific.

In the neural field model assumption *(i)* relates to the emission coefficients $\lambda(|\alpha - EL(r)|)$, assumption *(ii)* to the absorption coefficients $\sigma(|\alpha - EL(r)|)$. The emission distribution $\lambda(|\alpha - EL(r)|)$ decreases sharply for the excitatory and smoothly for the inhibitory connections with increasing difference between the emission direction $\alpha$ and the orientation of the elongation of the connection patterns $EL(r)$ (i.e. the preferred orientation). For excitatory connections the absorption and emission distributions are similar, whereas for inhibitory connections the absorption coefficients $\sigma(|\alpha - EL(r)|)$ grow with increasing $|\alpha - EL(r)|$. Note that the sharp spike emission distribution

excitation is blurred by the random diffusion process described by Eq. 3.5.

### 3.2.4 Synapses

A spike emitted by population $s'$ that travels in the neural neural field (according to Eq. 3.5) evokes a postsynaptic potential if the target population $s$ absorbs it. The number of absorbed spikes per neuron is

$$a_{s's}(\mathbf{r},t) = \frac{1}{\phi_s(\mathbf{r})} \cdot w_{s's} \cdot \int_0^{2\pi} \sigma_{s'}^{(|\alpha - \mathrm{EL}(r)|)} f_{s'}^{(\alpha)}(\mathbf{r},t)\, d\alpha\,, \tag{3.6}$$

where $\phi_s$ is the density of neurons at point $\mathbf{r}$, $w_{s's}$ is the connection probability ($\sum_s w_{s's} = 1$) between neurons of types $s'$ and $s$. The postsynaptic conductance is the convolution of the synaptic interactions $a_{s's}(\mathbf{r},t)$ with the alpha-function

$$\gamma_{s's}(\mathbf{r},t) = \frac{\bar{\gamma}_{s's}}{\tau_{s's}} \cdot \int_0^{\infty} a_{s's}(\mathbf{r},t-t') \cdot t' \cdot e^{1-t'/\tau_{s's}}\, dt'\,, \tag{3.7}$$

where $\bar{\gamma}_{s's}$ is the maximal conductance, and $\tau_{s's}$ is the time to peak. The postsynaptic current is determined based on the postsynaptic conductance (Eq. 3.4).

### 3.2.5 Discretization

The discretization of the state space is done such the that the first two moments of the distribution functions $f_s^{\alpha}$ and $g_s$ is invariant of scaling. This can be done e.g., by choosing the discrete time step $\Delta t$ to be the single scaling factor that determines the other discretization units ($\Delta x$, $\Delta u$) such that the diffusion rates $D_u$ and $D_r$ are constant for all $\Delta t$. The scaling rules are such that in the limit of $\Delta t \to 0$ the other discrete units converge also to zero, and the continuous model is obtained. The interested reader can find scaling rules and the proofs in (Barna et al., 1998; Grőbler et al., 1998). For the present simulations the discretization units are the following: $\Delta t = 0.08$ msec; $\Delta x = 0.1$ mm; $\Delta u = 1.4$ mV.

### 3.2.6 Interpretation of the populational activity

The neural field model describes the *statistical* behavior of neural populations, and not individual neurons. It follows, that the neural activity is described in terms of the probability of firing. The probability of firing in a 3 msec intervals divided by 3 msec is interpreted as the expected firing rate. The leaky integrate-and-fire dynamics (Eq. 3.4) of a single cell is calculated in parallel to the stochastic model for each discrete spatial element. These "representative" cells receive the same recurrent and feed-forward input as the cell population in the discrete spatial element. This way we have access to single cell behavior also. Note that the calculated membrane potential in time, is not exactly the average, but the deterministic version of the stochastic diffusion equation for the density function that describes the neurons (Eq. 3.1).

## 3.3 Computational results

In the following we investigate the emergent orientation selectivity map, and its relation to the assumed anisotropy of the recurrent connections. We show that the emergent orientation selectivity

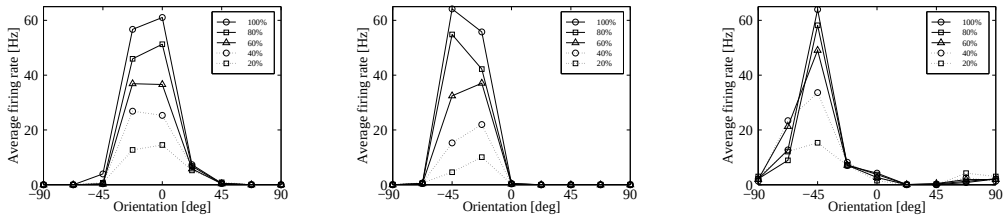


Figure 3.4: The orientation tuning curves of three discrete spatial elements at different Michelson contrast levels.

is contrast invariant. Then we compare the populational neural activity with the membrane potential of an "representative" single neuron in a discrete element. Our simulations reveal that although the cortical units receive a prominent none orientation specific excitation from the LGN, the firing activity of the most cells is highly orientation specific.

### 3.3.1 Emergent orientation selectivity

Figure 3.4 shows the orientation tuning curves of cell populations in three discrete spatial elements at different contrasts. The tuning curves were determined by presenting a series of eight gratings with different orientations. Here the time averaged expected firing rate of the excitatory cell population is plotted. Stimuli of 150 msec duration were interleaved by 15 msec blanks. The majority of the spatial elements with receptive fields overlapping the stimulus patch showed sharp contrast invariant tuning. The shape of the tuning curves is different for different locations because of the inherent scatter in the connection pattern. The initial elongation map (bars in Fig. 3.2) obtains only a smoothness constraint, otherwise it lacks regularities. The spike diffusion is also stochastic. The orientation tuning curves may also differ slightly for different contrasts because of the noise in the neural state dynamics Eq. 3.1.

For different locations $\mathbf{r}$ the emergent preferred orientation $\mathrm{PO}(\mathbf{r})$ (Fig. 3.5 right, bars) and orientation selectivity index $\mathrm{OSI}(\mathbf{r})$ (Fig. 3.5 right, shading) are determined from the average firing rate $\mathrm{resp}(\mathbf{r},\alpha_i)$ of model neurons for grating stimuli with orientations $\alpha_i = (i-1)\cdot 22.5°$. The preferred orientation (PO) is the circular mean of the spike distribution as a function of orientation. The orientation selectivity index (OSI) relates to the angular variance (Batschelet, 1981) of this distribution.

$$\mathrm{PO}(\mathbf{r}) = \arctan\left(\frac{\bar{y}}{\bar{x}}\right) \;; \quad \mathrm{OSI}(\mathbf{r}) = \frac{\sqrt{\bar{x}^2+\bar{y}^2}}{E_{\alpha_i}[\mathrm{resp}(\mathbf{r},\alpha_i)+a]}\,, \tag{3.8}$$

where $\bar{x} = \frac{1}{n}\sum_i \mathrm{resp}(\mathbf{r},\alpha_i)\cos(2\cdot\alpha_i)$ and $\bar{y} = \frac{1}{n}\sum_i \mathrm{resp}(\mathbf{r},\alpha_i)\sin(2\cdot\alpha_i)$, and $E_{\alpha_i}$ denotes the averaging over different orientations. In the definition of the orientation selectivity index $a = 2$ in the denominator relates to the significant activity level such that the OSI $\to 0$ if $E_{\alpha_i}[\mathrm{resp}(\mathbf{r},\alpha_i)] < a$. Our simulations reveal that the connection anisotropy $EL(r)$ and the emergent preferred orientation $PO(r)$ are strongly correlated (circular correlation coefficient = 0.86) (Fig. 3.6).

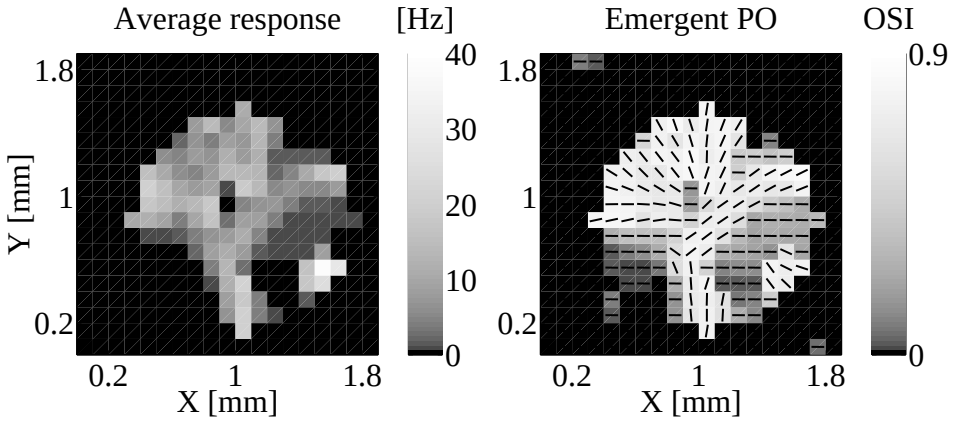


Figure 3.5: **Left:** The average cortical firing rate (shading) in the model as a response to sinusoidal grating patches with different orientations. **Right:** The emergent preferred orientations (bars) and the orientation selectivity index (shading). The preferred orientation is showed only for spatial elements with OSI(r)>0.5.

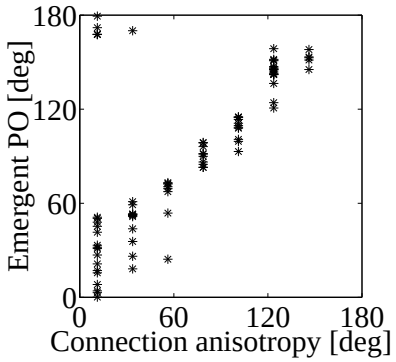


Figure 3.6: The relation between the assumed connection anisotropy $EL(r)$ and the emergent preferred orientation.

### 3.3.2 Orientation tuning dynamics

The firing dynamics of an "average" cell at 100% contrast (Fig. 3.7, top) show immediate tuned response without any spikes at off optimal or null orientations. The dynamics of the expected firing rate for a neural population in spatial element (Fig. 3.7, bottom) show that the initially poorly specific expected firing rate is around $0.1-2$ Hz. These simulation results indicate that non orientation-specific cortical activity with a very low probability is sufficient to establish sharply orientation tuned response via the specific recurrent amplification. In the model the dynamics of the emergence of orientation tuning are in accordance with experimental observations from a relatively small sample

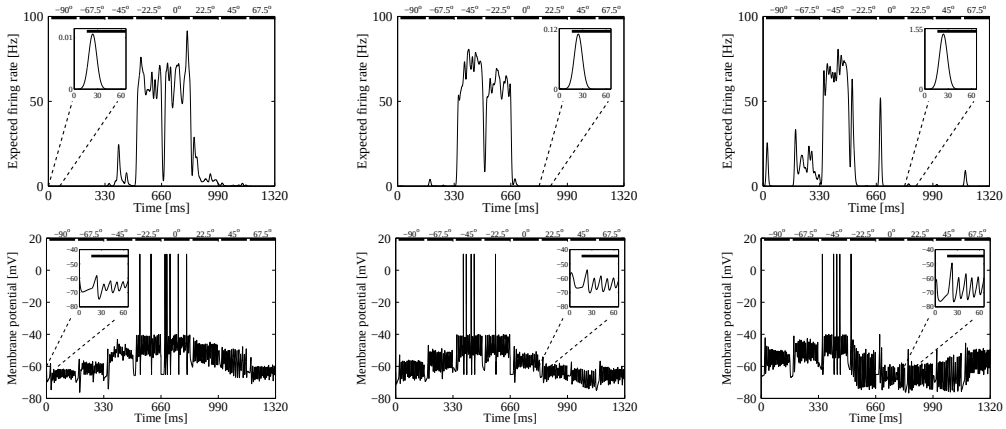


Figure 3.7: The dynamics of neural activity. The stimuli are 100% contrast sinusoidal gratings with different orientations as indicated at the top of the figure. The bars denote the stimulus onset, the gaps between the bars are blank periods. The orientation tuning curves in Fig. 3.4 are calculated for the same three locations. **Top** The dynamics of the expected firing rate. **Bottom** The membrane potential dynamics of the "representative" neurons from the three locations.

of neurons

(Celebrini et al., 1993; Pernberg et al., 1998) which show immediate emergence of tuned response and with experiments which use a large sample of observations to detect the small initial non-specific response (Ringach et al., 1997*a*). In contrast to the firing dynamics, the membrane potential dynamics show a pronounced initial non-specific EPSP (Fig. 3a, inlet) similar to observations for cat (Pei et al., 1994; Pernberg et al., 1998).

## 3.4 Conclusions

In a neural field modeling framework we further tested our "intracortical hypothesis" introduced in the previous chapter and in (Adorján, Levitt, Lund and Obermayer, 1999). The investigation focused on *(i)* the relation between the geometry of the recurrent connectivity pattern and the emergent orientation selective simple cell response; *(ii)* the orientation tuning dynamics.

Here we extended the original neural field model (Grőbler et al., 1998; Barna et al., 1998) that it includes specific recurrent connections. In the model definition, we used the immediate consequence of the assumption for the intracortical hypothesis that recurrent excitatory connections are anisotropic with an elongation parallel to the preferred orientation. From the connection anisotropy it follows that from the drifting direction of a spike the type of emitting and the target neuron can be inferred.

The neural field modeling framework allowed us to investigate in good spatial resolution the emergent orientation maps in relation to a *geometrically detailed* and realistic recurrent connectivity. In the model only a moderate connection anisotropy is assumed (cf. Fig. 3.2). Results show that this small anisotropy can provide a strong enough initial orientation bias even in a noisy environment. It follows that a possible experimental test has to take into account the bouton density maps of cortical

neurons with a great precision in the visual field coordinate system. This latter requirement may be an important if there is a small scatter in the retinotopy. Comparing the first principle component of the synapse distribution with the preferred orientation of the injected cortical cell would be an exact and easy method to apply.

Furthermore, we investigated the *orientation tuning dynamics* of populational spiking activity and the membrane potential of single cells. Our study reveals that the gradual sharpening of the orientation tuning of the membrane potential and the immediate sharp tuning of the spiking activity of single cells do not contradict to each other. In the model the non-specific excitation arriving from the LGN causes the initial unspecificity of the membrane potential orientation tuning. In most cases this unspecific excitation remains below threshold. However, random fluctuations may drive a small portion of neurons beyond threshold. The probability of this unspecific initial firing of a group of neurons with a certain connection anisotropy depends on the stimulus orientation and constitutes the initial orientation bias. This bias is then further amplified and sharpened by the specific recurrent excitatory connections and by the recurrent inhibition.

In conclusion, we have shown in biologically plausible model that incorporates noise and irregularities in the recurrent connections that it is possible to generate orientation selective response purely intracortically. The simulations offer a possible explanation for the experimental data regarding the orientation tuning dynamics of cortical cells in V1. These experimental results would be hard to explain assuming orientation bias originating from convergence pattern of geniculate cells to cortical cells, as it was assumed in the previous works (Ben-Yishai et al., 1995; Douglas et al., 1995; Somers et al., 1995).

# Chapter 4

# Dynamic coding: from the salient towards the details


## Abstract

The role of recurrent cortical amplification in the generation of orientation selective response in the primary visual cortex (V1) is still a controversial issue. Model studies have shown that strong recurrent cortical interactions support sharp contrast invariant orientation tuning, but have the disadvantage that complex stimuli cannot be represented reliably in the activity pattern across a cortical hypercolumn. Weak recurrent cortical interactions—on the other hand—allow for the reliable representation of complex stimuli, but do not lead to the sharp and contrast invariant orientation tuning which is observed experimentally.

Here we propose that these problems can be resolved if the strength of recurrent cortical interactions is dynamic and changes on the time scale of a typical eye fixation period. In our *"dynamic cortical amplifier"* model, the recurrent cortical competition is strong in the initial phase signaling the salient orientation in a winner-takes-all fashion. In the *second phase*, when the strength of recurrent connections decreases, finer details of a complex stimulus can be represented reliably. This provides a compromise between the "feed-forward" and the "recurrent" hypotheses for orientation selectivity. The model predicts that the signaling of a lower intensity edge is delayed if it is masked by an edge with higher intensity.

We then show that decreasing the recurrent competition with time naturally follows from functional considerations, i.e. from the requirement that the mutual information between stimuli and representations is maximal for any time interval beginning with stimulus onset. The key observations are that the signal-to-noise ratio of cortical representations increases with time (because more spikes are available) and that the optimal strength of the recurrent connections (w.r.t. information transfer) decreases with output noise. Consequently the model predicts that the information content per spike (or the SNR for a fixed time window) decreases with time for a flashed static stimulus in accordance with recent experimental studies. The neural system thus adapts to its own internal changes by modifying its coding strategy, a phenomenon which one may refer to as *"dynamic coding"*.[1]


[1]This chapter is based on (Adorján et al., 2000). Special thanks to Lars Schwabe who provided the results for the detailed computational model. The idea of information maximization in increasing time windows was inspired by the several discussions with Christian Piepenbrock.

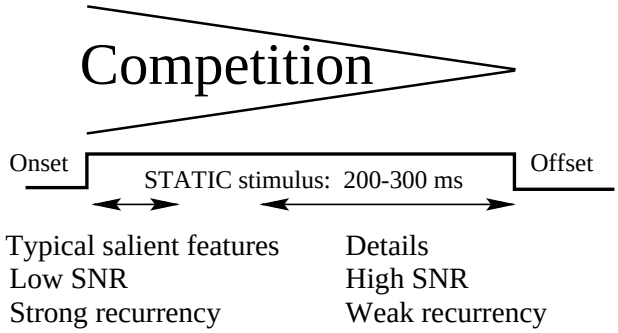


Figure 4.1: The model hypothesis. After stimulus onset (at the beginning of a fixation period) the recurrent competition is high and it attenuates on a short time scale. The fast modulation of the recurrent competition can be motivated by the necessity to be able to represent multiple features, and at the same time establish a robust and contrast invariant sharpening of the feed-forward signal. We propose that the contrast invariant sharpening is established in the first phase of the cortical processing in which the recurrent competition is strong. In the second phase, weakening of the recurrent connection strength allows the representation of multiple orientations in the second phase of the response. Decreasing the recurrent cortical competition follows naturally from functional considerations, i.e. from the requirement that the mutual information between stimuli and representations is maximal for any time interval beginning with stimulus onset. We argue, that the signal-to-noise ratio (SNR) changes as the time window for the representation increases and that the optimal strength of the recurrent connections (w.r.t. information transfer) decreases with output noise.

## 4.1 Introduction

In the last four decades there has been a vivid and highly polarized discussion about the role of the recurrent competition in the primary visual cortex (V1) (see (Sompolinsky and Shapley, 1997) for review). The main question is whether the recurrent excitation sharpens a weakly orientation tuned feed-forward input, or the feed-forward input is already sharply tuned, hence the massive recurrent circuitry has a different function. Strong cortical recurrency implements a highly nonlinear mapping of the feed-forward input and obtains robust, sharply tuned cortical response even if only weak or no feed-forward orientation bias is present (Ben-Yishai et al., 1995; Somers et al., 1995; Adorján, Levitt, Lund and Obermayer, 1999). However, such a competitive network in most cases fails to process multiple orientations within the classical receptive field and may signal spurious orientations (Carandini and Ringach, 1997). Also, strong local competition can prevent long range modulation of the tuning curve *shapes* from outside the classical receptive field. This motivates the concept that the primary visual cortex should rather map its input in a less competitive (more linear) fashion, but requires sharply orientation tuned feed-forward input to the primary visual cortex (Hubel and Wiesel, 1962; Troyer et al., 1998).

Although previous models for orientation selectivity in V1 vary on a wide scale, they have one common feature: each of them assumes that the synaptic strength and the excitability of the cortical cells is constant on the time scale on which the network operates, that is during a fixation period (200 – 300 msec). Given the phenomenon of fast synaptic dynamics and slow intracellular adaptation this, however, does not need to be the case. Here we show, that short term synaptic dynamics of the

recurrent excitatory synapses allow a cortical network to operate in both—competitive and linear—regimes. It is likely that decreasing recurrent competition can be implemented also by intracellular adaptation mechanisms, however this hypothesis has to be further investigated.

The major hypothesis of the current study is that the recurrent competition attenuates gradually after stimulus onset (see Fig. 4.1). We make the hypothesis that such a fast modulation of the recurrent cortical competition is advantageous for two reasons. *(i)* The network is capable to sharpen the feed-forward signal and also able to represent multiple features. *(ii)* In contrast to all previous studies that tried to assemble the *neurophysiological* "puzzle" into a coherent image and aimed to understand cortical dynamics from a mechanistic point of view, we extend our view to *functional* aspects of short term dynamics of the visual cortical response. We propose that decreasing recurrent competition obtains maximal information transfer in any time interval beginning with the stimulus onset. In other words, decreasing the initially strong recurrent competition naturally follows from functional considerations.

Firstly, we show in a detailed bottom-up computational framework (section 4.2) that such a network can establish sharp orientation tuning from a broadly tuned feed-forward input, while it is still capable to respond correctly to multiple orientations. We compare the response of a detailed cortical amplifier model with and without fast synaptic plasticity at the recurrent excitatory connections to single and multiple bars within the classical receptive field. The model makes testable predictions about the time course of representation of complex stimuli.

Secondly, we relate this dynamical property to function by considering the transmission of information with noisy cortical units about a *static* input (section 4.3). We find that if the neural activation function changes with time from a highly competitive to a more linear mode, then the information transfer in any increasing time interval after stimulus onset is largely increased compared to the information transfer achieved by a static neural activation function. The fast modulation of the recurrent connection strength is motivated by fast changes of the internal state of the cortical network and not by the stimulus dynamics. Fast modulation of the coding strategy establishes a *dynamic code*.

## 4.2 Dynamic cortical amplifier

### 4.2.1 The model setup

To investigate our first hypothesis, we set up a model for an orientation-hypercolumn in the primary visual cortex with similar structure and parameters as in (Carandini and Ringach, 1997). The important novel feature of our model is that fast synaptic depression is present at the recurrent excitatory connections. Neurons in the cortical layer receive orientation tuned feed-forward input from the LGN. Cortical neurons are, furthermore, connected to each other via highly orientation specific excitatory and non-specific inhibitory connections. The resulting recurrent kernel has a shape similar to a Mexican-hat in orientation space. Cortical cells tuned to similar orientations excite each other and effectively inhibit cells tuned to the orthogonal orientation. The recurrent and feed-forward excitatory synapses exhibit fast depression due to the activity dependent depletion of the synaptic transmitter (Abbott et al., 1997; Tsodyks and Markram, 1997). The model structure is depicted in Fig. 4.2.

The membrane potential $V(\theta,t)$ (Eq. 4.1) of a cortical cell tuned to an orientation $\theta$ decreases due to leakage and recurrent inhibition $I^{\text{inh}}$, and increases due to the recurrent excitation $I^{\text{exc}}$ and the

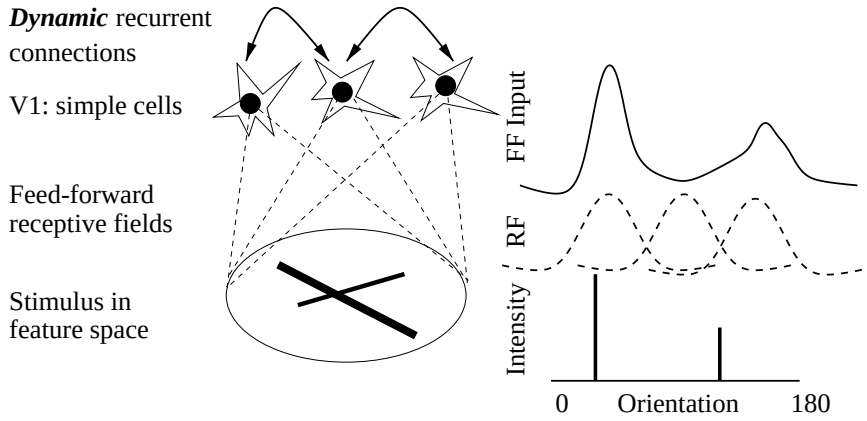


Figure 4.2: The structure of the model. A visual stimulus is represented in the orientation (feature) space. This abstract stimulus representation allows us to investigate cortical feature representation on a general level and to avoid making strong assumptions about the early visual processing which is not a subject of the current study. The "intensity" of an orientation can be interpreted as contrast or other properties, like spatial frequency that modify the strength of the geniculate response. The feed-forward input from the LGN to the cortical simple cells is calculated by convolving the stimulus with the broad orientation tuning curve of the summed geniculate input to a cortical cell. (Note that this figure is only a sketch.) The feed-forward signal is re-amplified via the dynamic recurrent connections.

geniculate input $I^{\mathrm{LGN}}$

$$\tau_{\mathrm{m}} \frac{\partial}{\partial t} V(\theta,t) + V(\theta,t) \quad = \quad I^{\mathrm{LGN}}(\theta,t) + I^{\mathrm{exc}}(\theta,t) - I^{\mathrm{inh}}(\theta,t), \tag{4.1}$$

$\tau_{\mathrm{m}}$ (15 ms, McCormick et al. (1985), Stratford et al. (1996)) is the membrane time constant. The recurrent excitatory and inhibitory cortical inputs are given by

$$I^{\alpha}(\theta,t) \quad = \quad \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} J^{\alpha}(\theta,\theta',t) \exp\left(-\frac{\Delta(\theta',\theta)^2}{2\sigma_{\alpha}^2}\right) (f(\theta',t)+\eta)\, d\theta' \tag{4.2}$$

where $\Delta(\theta',\theta)$ is a $\pi$ periodic circular difference between the preferred orientations, $J^{\alpha}(\theta,\theta',t)$ are the excitatory and inhibitory connection strengths (with $\alpha \in \{\mathrm{exc},\mathrm{inh}\}$, and $f$ is the presynaptic firing rate. The maximal excitatory and inhibitory synaptic efficacies ($J^{\mathrm{exc}}_{\mathrm{max}} = 0.2$ mV/Hz and $J^{\mathrm{inh}}_{\mathrm{max}} = 0.8$ mV/Hz) are determined such that they remain in a biologically realistic regime and the network operates in a highly competitive regime after stimulus onset and obtains a contrast invariant orientation tuning. The excitatory synaptic efficacy $J^{\mathrm{exc}}$ is time dependent due to the fast synaptic depression, while the efficacy of inhibitory synapses $J^{\mathrm{inh}}$ is assumed to be constant. The recurrent excitation is sharply tuned $\sigma_{\mathrm{exc}} = 7.5°$, while the inhibition has broad tuning $\sigma_{\mathrm{inh}} = 90°$. The recurrent connectivity structure is determined in accordance with neurophysiological and anatomical data (for review see chapter 2 and, e.g., Sompolinsky and Shapley (1997)). The mapping from the membrane potential to firing rate is approximated by a linear function with a threshold at 0

| | |
|---|---|
| Membrane time constant $\tau_m$ (Stratford et al., 1996) | 15 ms |
| Recurrent excitatory strength $J_{max}^{exc}$ | 0.2 mV/Hz |
| Recurrent inhibitory strength $J_{max}^{inh}$ | 0.8 mV/Hz |
| Recurrent excitatory width $\sigma_{exc}$ | 7.5° |
| Recurrent inhibitory width $\sigma_{inh}$ | 90° |
| Activity gain $\beta$ | 0.13 Hz/mV |
| Feed-forward orientation tuning width $\sigma_{LGN}$ | 18° |
| Feed-forward synaptic recovery time constant $\tau_{rec}^{LGN}$ | 120 ms |
| Cortical synaptic recovery time constant $\tau_{rec}^{Ctx}$ | 850 ms |
| Feed-forward transmitter release probability $p^{LGN}$ | 0.35 |
| Cortical transmitter release probability $p^{Ctx}$ | 0.55 |

Table 4.1: Summary of model parameters used in this study.

($f(\theta) = \beta \max(0, V(\theta))$, $\beta = 15$ Hz/mV). Gaussian-noise $\eta$ with variances of 6 Hz and 1.6 Hz is added to the input intensities and to the output of cortical neurons respectively.

The orientation tuning curves of the feed-forward input $I^{LGN}$ are Gaussians ($\sigma_{LGN} = 18°$) resting on a strong additive orientation independent component. This feed-forward orientation tuning would correspond to a geniculo-cortical connectivity pattern with an approximate aspect ratio of 1:2. The orientation independent component is included to account for the non-zero volume of the feed-forward receptive field profile that is reflected in the strong geniculate background (or DC) activity (Kaplan et al., 1987). Both, the orientation dependent and independent components increase with contrast.

Considering a free-viewing scenario where the environment is scanned by saccading around and fixating for short periods of $200 - 300$ ms we model stationary stimuli present for 300 ms. The stimuli are one or more peaks in the feature space. These peaks correspond to overlayed bars or gratings with different orientations in real space. The exact translation from the abstract feature space to the geometrical arrangement of bars or gratings in a real space depends on the exact model for the retinal and geniculate receptive field profiles and the activation functions.

Note that in the current study we want to demonstrate that fast dynamics of cortical activity can be explained purely by the non-stationarity of the noise on the cortical code. Therefore here we do not consider factors which make the feed-forward input dynamic on a short time scale, like dynamic visual stimuli, micro-saccades, ocular drift and tremor.

Feed-forward and recurrent excitatory synapses exhibit fast depression. Fast synaptic depression is modeled by the dynamics of the expected synaptic transmitter or "resource" $\bar{R}(t)$ for each synapse. The amount of the available transmitter decreases proportionally to the transmitter release probability $p$ and to the presynaptic firing rate $f$ (see appendix C.2), and it recovers exponentially ($\tau_{rec}^{LGN} = 120$ ms, $\tau_{rec}^{Ctx} = 850$ ms, $p^{LGN} = 0.35$ and $p^{Ctx} = 0.55$),

$$\frac{d}{dt}\bar{R}(t) \quad = \quad \frac{1-\bar{R}(t)}{\tau_{rec}} - f(t)p(t)\bar{R}(t) \; = -\frac{\bar{R}(t)}{\tau_{eff}(f(t),p(t))} + \frac{1}{\tau_{rec}}. \tag{4.3}$$

The change of the membrane potential on the postsynaptic cell at time $t$ is proportional to the released transmitter $pR(t)$. The transmitter release probabilities and the recovery time constants are chosen from an experimentally observed range (Abbott et al., 1997; Tsodyks and Markram, 1997). The role

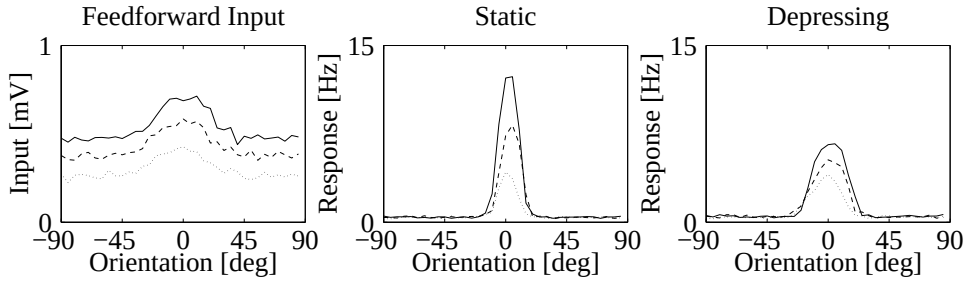


Figure 4.3: The feed-forward input (a), the response of the cortical amplifier model with static recurrent synaptic strength (b) and a network with fast synaptic depression (c) if the stimulus is single bar with different stimulus intensities. The cortical response is averaged over the first 100 ms after stimulus onset.

of the cortical synaptic parameters are further discussed in section 4.2.2. The excitatory connectivity strength between neurons tuned to orientations $\theta$ and $\theta'$ is expressed as $J^{\text{exc}}(\theta,\theta',t) = J^{\text{exc}}_{\max} pR_{\theta\theta'}(t)$. Similarly this applies to the feed-forward synapses. In the present model fast adaptation of the geniculo-cortical synapses accounts for the transient nature of the geniculate input. The exponential decrease in the feed-forward input defines a time window for the cortical processing in which the cortex receives strong feed-forward input and it prevents the emergence of oscillations in the cortical response. Note that the strength of the recurrent competition is independent of the strength of the feed-forward input, therefore fast adaptation at the recurrent excitatory synapses cannot be substituted by fast adaptation at the geniculo-cortical synapses. Fast synaptic plasticity at the feed-forward synapses has been investigated in more detail in previous studies (Artun et al., 1998; Adorján, Piepenbrock and Obermayer, 1999), see also chapter 5.

### 4.2.2 Computational results

In the following, we compare the predictions of the dynamic cortical amplifier model with fast synaptic depression at the recurrent excitatory connections with the classical static cortical amplifier model. In both cases fast synaptic depression is present at the feed-forward connections limiting the duration of the effective feed-forward input to $200 - 400$ ms. Figure 4.3 shows the orientation tuning curves at different stimulus intensities. The feed-forward input is noisy and broadly tuned (Fig. 4.3a). Both models exhibit contrast invariant tuning (Fig. 4.3b, c). The half-width at half-height of the cortical tuning curves is basically independent of the stimulus intensity, while it increases strongly for the feed-forward input. If fast synaptic depression is present at the recurrent excitation, the cortical network sharpens the broadly tuned feed-forward input in the initial response phase. Once sharply tuned response is established, the tuning width does not change, only the response amplitude decreases in time.

The predictions of the two models differ substantially if multiple orientations are present (Fig. 4.4). First, we test the cortical response to two peaks in feature space separated by 60° with different intensities (Figs. 4.4a, b). If the recurrent synaptic weights are static and strong enough (Fig. 4.4a), then only one orientation is signaled. The cortical network selects the feed-forward input with the

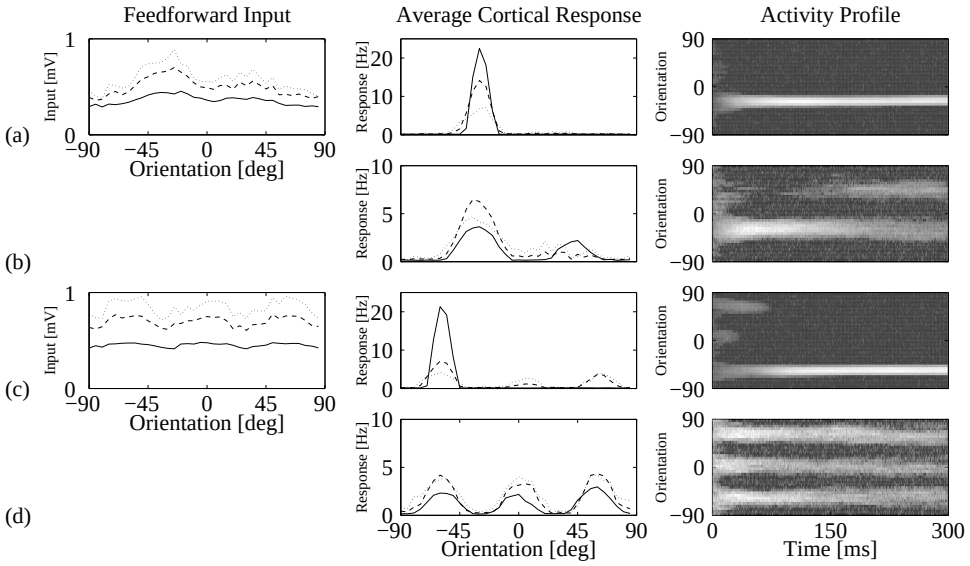


Figure 4.4: The response of the cortical amplifier model with static (a,c) and fast depressing recurrent synapses (b, d). In both models the feed-forward synapses are fast depressing. In the left column the feed-forward input is shown, that is same for both models. Two types of stimuli were applied. The first stimulus consists of a stronger ($\alpha = -30°$) and a weaker peak ($\alpha = +30°$) in feature space (a, b). The second stimulus consists of three equal intensity peaks with orientations that are separated by 60° (c, d). In the middle column the cortical response is shown averaged for different time windows ([0..30 ms] dotted; [0..80 ms] dashed; [200..300 ms] solid line). In the right column the cortical activity profile is plotted as a function of time. Gray values indicate the activity with bright denoting high activities.

highest amplitude in a winner-takes-all fashion. In contrast, if synaptic depression is present at the recurrent excitatory synapses, both peaks in feature space are signaled in parallel (at low release probability, Fig. 4.4b). First, those cells fire which are tuned to the orientation of the peak with the stronger intensity and a sharply tuned response emerges at a single orientation—the network operates in a winner-takes-all regime. The synapses of these highly active cells then become strongly depressed and cortical competition decreases. As the network is shifted to a more linear operation regime, the second orientation is signaled too. Note that this phenomenon—together with the observed contrast invariant tuning—cannot be reproduced by decreasing the static synaptic weights in the cortical amplifier model. In our dynamic cortical amplifier model the recurrent synaptic efficacy changes inhomogeneously in the network depending on the activity. Only the synapses of the highly active cells depress strongly and therefore sharply tuned response can be evoked by a peak with weak intensity. Fast synaptic depression thus behaves as a local self-regulation that modulates competition with a certain delay. This delay and therefore the delay of the rise of the response to the second peak depends on the effective time constant $\tau_{\text{eff}}(f(t), p) = \tau_{\text{rec}}/(1 + pf(t)\tau_{\text{rec}})$ of the synaptic depression at the recurrent connections. If the depression becomes faster due to an increase in the release probability $p$, then the delay decreases. The delay also scales with the difference between the intensities

of the two peaks. The closer to each other they are, the shorter the delay will be (compare Figs. 4.4b, d right column). For further detailed explorations see (Schwabe, 1999).

In Figs. 4.4c, d cortical response to three peaks in feature space with equal intensities is presented. In this case cells tuned to the presented three orientations respond in parallel if fast synaptic depression at the recurrent excitation is present (Figs. 4.4d). The cortical network with strong static recurrent synapses again fails to signal faithfully its feed-forward input. Additive noise on the feed-forward input introduces a slight symmetry breaking and the network with static recurrent weights responds strongly at the orientation of only one of the presented peaks (Fig. 4.4c).

In summary, our simulations revealed that a recurrent network with fast synaptic depression is capable to obtain robust sharpening of its feed-forward input and it also responses correctly to multiple orientations. A clear experimentally testable prediction of the model is that the response to a flashed bar with lower contrast can be delayed by masking it with a second bar with higher contrast (Fig. 4.4b, right). We also suggest that long range integration from outside of the classical receptive field could emerge in a similar fashion. In the initial phase of the cortical response strong local features are amplified. In the second longer phase, the recurrent cortical competition decreases, and the long-range modulatory recurrent and the feed-forward input has a stronger effect compared to the local recurrent excitation. In the following, we investigate whether this strategy is favorable in terms of information transfer in time.

## 4.3 Dynamic code

In the previous section we investigated the role of recurrent cortical amplification in a detailed computational model for a cortical hypercolumn. We concluded that both, contrast invariant orientation tuning and representation of multiple orientations can be achieved by a recurrent cortical network if the strength of the recurrent amplification attenuates after the beginning of a fixation period in a free-viewing scenario or alternatively after stimulus onset. We suggested and explored fast synaptic depression as a plausible neurophysiological mechanism that may implement the fast weakening of the recurrent competition between active neurons. Alternatively, intracellular adaptation of excitatory neurons could also be considered (Schwabe et al., 2000). New experiments and computational studies could help to decide between these possibilities.

In the present section the role of recurrent cortical competition is investigated from the *functional* point of view. Here we also consider the natural situation for our visual system: in a free-viewing scenario the visual world is explored by fixating at different positions by fast saccades. Fixation periods interleaved with saccades result in a series of visual stimuli that are flashed onto the retina for the length of a fixation period ($\Delta T = 200 - 300$ msec). Restricting our study for the representation of *static* stimuli, we aim to prove that under certain conditions, weakening the initially strong competition between orientation selective cells (feature detectors) is optimal in the terms of information transmission in time.

The key observation is the following. The internal state of the nervous system changes in time and the coding strategy may adapt to the internal changes in order to fit to some optimality criteria (that may also change): the neural code is dynamic. Therefore neural adaptation is not necessarily driven only by environmental alterations, but also may be a consequence of alterations in the *internal* conditions. If the coding strategy changes fast, this may have the consequence that the fine temporal structure of the spike pattern becomes relevant for coding. This gives a relation of the here introduced "dynamic code" to the established concept of "temporal code".

Specifically, here an important aspect of the internal neural state, the *noise* level on the representing cortical neurons is considered. We assume that the entire activity pattern from stimulus onset until time $t$ constitutes the neural code about a static stimulus. The spike timing has a limited precision (e.g., Holt et al., 1996), but the capacity of the encoding neurons increases with time because as the duration of the coding interval increases, more spikes are available for representation. In other words, the noise on the neural code (the output noise) is *non-stationary*. Note that this is entirely independent of the actual neural code and therefore it is true for rate code, as well as for fine time-scale temporal code. This dynamic nature of the output noise was neglected in previous studies. Here we demonstrate that if the time interval $\Delta T$ that is available for encoding is larger than the average inter-spike-interval, then the non-stationarity of the output noise has important implications on neural coding and the short term dynamics of neural activity.

We assume that in the chain of the encoding and read-out mechanisms, the read-out mechanism has a memory that is long enough for taking into account several input spikes. This requires the presence of long-term ($\Delta T >$average inter-spike-interval) integrative mechanisms in the cortical circuits.[2] After stimulus onset a larger and larger time window is available for representation, consequently the signal-to-noise ratio (SNR) increases. In other words, at the very beginning of a fixation period only a few spikes are accessible for encoding, the uncertainty about the stimulus is high. Towards the end of a fixation period the increasing number of spikes can give a representation with a higher fidelity. Note again that the SNR increases simply because the average spike count increases with time, and—independently of the neural code!—more spikes are available for estimating the message encoded by the neurons.

As the noise on the cortical code changes with time during a fixation period, the optimal way of processing the feed-forward visual signal also changes. Using the information transfer as a quality measure and given the above assumptions here we address the following question: *What kind of input-output mapping dynamics maximize the information transfer in any increasing time interval beginning with the stimulus onset, given a signal distribution and the output noise dynamics?* Optimal information transfer in any increasing time interval beginning with the stimulus onset has a clear evolutionary advantage. For instance reaction times shorter than a fixation period can save life in some situations. This opens us a space for speculation that the visual system of evolutionary advanced species (for which visual signals are important) is also close to be optimal in the above sense. We proceed with an information theoretical investigation that gives predictions and highlights possible design principles that determine the short-term dynamics of visual cortical processing.

Let us sum up our conceptually essential assumptions and proposals:

- The entire activity pattern within a period that is longer than the average inter-spike-interval constitutes the neural code about a stimulus within a fixation period. It follows that the signal-to-noise-ratio increases significantly with time because more spikes are available for representation.
- The optimality criteria for the neural code is the information transfer in any increasing time interval beginning with the stimulus onset.
- The stimulus is static. It follows that we can study short-term neural dynamics that is purely a consequence of internal changes. (Neural dynamics as a consequence of changing stimulus have been investigated by numerous previous works.)

[2] Quantitative experimental and theoretical studies are still have to be done for visual cortical neurons to justify or falsify this essential assumption.

On our way of estimating the optimal dynamics of the cortical transfer function we make some technical simplifications. Firstly, we constrain the possible mapping functions to a family of functions parameterized with a single parameter. Namely, we model the competitive visual cortical mapping by the "soft-max function". The inverse temperature parameter $\beta$ can be interpreted as the strength of recurrent competition. The particular choice of this function family was motivated by our original question: How strong is the recurrent cortical competition in V1? The soft-max function with high competition parameter well approximates a cortical amplifier circuitry, while it also accounts for the pure feed-forward scenario if the competition parameter is low. The inverse temperature $\beta$ has been interpreted similarly as a parameter for the recurrent competition strength in a previous study (Piepenbrock and Obermayer, 1999). Secondly, the optimal dynamics of the recurrent competition level $\beta(t)$ is approximated here by the a competition level that is optimal for an output noise variance at time $t$ as if the competition was static in the interval of $[0..t]$. A future study could give a more exact approximation for the competition dynamics.

In the following, we describe and motivate the abstract soft-max model for the recurrent cortical network. For the soft-max model we study the information transfer dynamics for two different output noise models. Firstly, the noise on the neural code is modeled as additive Gaussian on an estimate for the firing rate. Secondly, the encoding neurons are modeled as Poisson-firing units. This latter is a more realistic approach.

### 4.3.1 The abstract model

Instead of the detailed, bottom-up model for the recurrent cortical network in V1, here we use a simplified abstract framework. The simplification was necessary for two reasons. First, the estimation of information transfer requires to obtain a large amount of input/output samples that is computationally demanding and cannot be easily produced by a detailed numerical model. Second, it was necessary to formalize the visual cortical mapping such that the strength of the recurrent competition can be related to some simple parameters in the model. It follows, that we need a mapping function that accounts well for the function of the recurrent network of an orientation hypercolumn in V1, it is easy to compute and it has a parameter that clearly determines the level of the recurrent competition. Switching this competition parameter from low towards high values provides the transition from the purely feed-forward to the recurrent processing of the feed-forward input.

The cortical activity $\vec{y}$ as a function of the input $\vec{x}$ is modeled here as

$$\vec{y}(t) = g(\vec{x},t) + n(g(\vec{x},t),t) = \vec{u}(t) + n(\vec{u}(t),t) \;, \tag{4.4}$$

where the mapping $g : X \to Y$ is deterministic and $n(\vec{u}(t),t)$ is an additive or multiplicative noise process. The static input $\vec{x}$ represents the feature intensities and it is interpreted here as the feed-forward input to cortical neurons in orientation space, thus $x_i$ is the intensity at orientation $i$. The mapping $g(\vec{x},t)$ refers to the transformation of the feed-forward input $\vec{x}$ by the recurrent network, $\vec{u} = g(\vec{x},t)$ is the expectation value for the neural activity and $n(\vec{u}(t),t)$ is the inherent noise on the cortical neurons' activity. The input-output mapping and the noise are functions of time $t$ with $t = 0$ at the stimulus onset.

We model the input-output mapping by the "soft-max function"

$$g_{\mathrm{i}}(\vec{x},t) = \mu(t)\frac{\exp(\beta(t)x_i)}{\sum_i \exp(\beta(t)x_i)} \;. \tag{4.5}$$

Cortical units with different preferred orientations but with overlapping receptive fields are indexed by $i$. The expected spike count in the entire network is $\mu(t)$. The $\beta$ parameter can be interpreted as the level of recurrent competition. As $\beta \to 0$ the network operates in a more linear mode, while a highly nonlinear winner-takes-all mode arises if $\beta \to \infty$. In all cases the average activity in the network is constrained that has been suggested to minimize metabolic costs (Baddeley, 1996). Furthermore, it has been already argued that the recurrent network with orientation specific connections implements competition between the feature detectors via a divisive operation (e.g., Carandini, Heeger and Movshon, 1997; Lee et al., 1999). Thus the soft-max function fits to our initial requirements.

For the sake of simplicity here we assume that the neural code is the spike count in the increasing time intervals after stimulus onset. Note, however, that our results rely purely on the fact, that noise on the neural code decreases as the time window for encoding increases because more spikes are available for representation. This is true for any neural code that is formed by spike patterns in increasing time windows, only the time course of the noise depends on the particular alphabet that is formed by the spike patterns.

We investigate two noise models. First, the output $\vec{u}$ is interpreted as the firing *rate* that sums up to $\mu(t) = 1$ independent of the time after stimulus onset. In that case $n(\vec{u}(t), t) = n(t)$ represents the noise on the rate estimation and it is assumed to be Gaussian with a standard deviation of $\sigma(t)$. As the time window with a width of $t$ for representation increases, $\sigma(t)$ decreases by a factor that is proportional to the square root of $t$ because the estimate for the spike count can be based on more samples. Second, a more realistic model is studied. The cortical spiking is assumed to follow Poisson statistics with a mean of $g_i(\vec{x}, t)$. The non-stationarity of the noise arises from the fact that the mean of the summed spike count $\mu(t)$ in the entire network increases with time proportionally to the instantaneous firing rate.

To estimate the information transfer, we need to define a model for the input distribution. If the visual processing is adapted to process optimally our visual world, we need to define a realistic model for the stimulus distribution in order to reach proper conclusions. Edges are shown to be largely independent components of the natural scenes (Olshausen and Field, 1996; Bell and Sejnowski, 1997). The input to the orientation selective visual cortical cells depends on the orientation of the edge in the receptive field. It follows, that a factorizing distribution is a realistic model for the input. The prior distribution of the different orientation intensities is assumed to be

$$p(\vec{x}) = \frac{1}{Z}\Pi_i \exp\left(\frac{-x_i^\alpha}{\xi}\right) \quad \text{for } x \geq 0\,, \tag{4.6}$$

where the exponent $\alpha$ determines the sparsity of the probability density function, $Z$ is a normalizing constant and $\xi$ determines the variance. For different $\alpha$ values, the variance is kept constant by the proper choice of $\xi$. If $\alpha = 2$, the input density is the positive half of a multivariate Gaussian distribution. With $\alpha > 2$ the signal distribution becomes sub-Gaussian and with $\alpha < 2$ it becomes super-Gaussian (sparse). An experimental study indicates that edges in natural scenes form a sparse distribution (Ruderman, 1994), therefore the our stimulus distribution model with $\alpha < 2$ mimics the best the real world situation.

The input distribution together with the cortical mapping and the noise model constitutes the framework to model cortical mapping and estimate information transfer. In the following we present calculations for a cortical mapping model, where the both, the input and output spaces are three-dimensional ($D = 3$). (Note that the complexity of the mutual information estimation increases exponentially with the number of dimensions.)

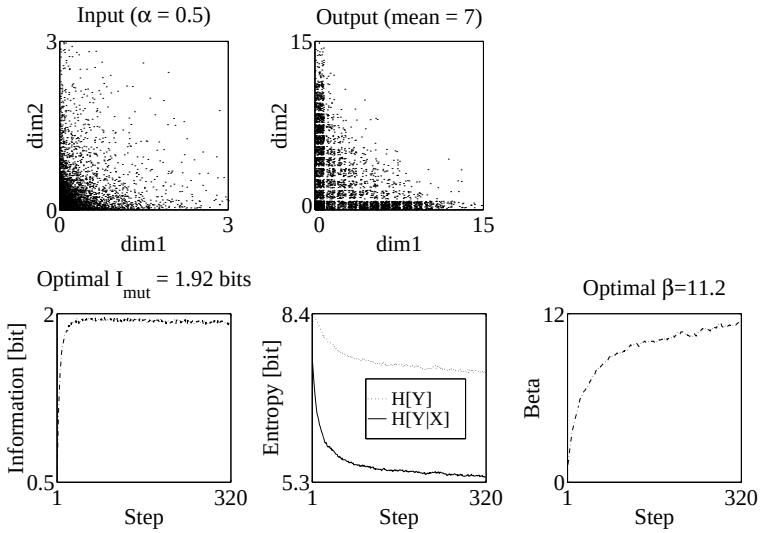


Figure 4.5: Estimation of the optimal competition parameter β if the input and the output space are three-dimensional. The output units fire according to a Poisson process with a mean spike count of 7. For this particular example the stimulus distribution is sparse $\alpha = 0.5$. **Top row:** The two-dimensional projection of the three-dimensional input and output distribution. The output distribution is illustrated here only for the optimal competition level: $\beta = 11.2$. The dimensions refer to different features/feature detectors. Each dot in the input distribution represents an input with given feature intensities. The input can interpreted as the feed-forward representation of a part of an image that falls into the cortical neurons' receptive fields. The spike count of the output neurons (different feature detectors) is also represented by a dot that falls into a little square that is centered at the observed spike count. (Note that the dots are jittered around each spike count coordinate only for demonstration reasons.) **Bottom row:** The mutual information ($I[X,Y] = H[Y] - H[Y|X]$); the output entropy $H[Y]$ (dotted line) and the conditional uncertainty $H[Y|X]$ (solid line); the competition parameter β as a function of learning steps. Note that the β does not saturate as a function of the learning step because we increase the learning rate on β exponentially. This speeds up the convergence of the mutual information. The incorporation of this improvement into our learning algorithm was motivated by our preliminary calculations that showed long flat plateau of mutual information as a function of β. The learning procedure is described in detail in appendix B.

### 4.3.2 Optimizing the recurrent competition dynamics

We approximate the optimal transfer function dynamics as follows: *(i)* We consider only a function family the soft-max function that accounts well for cortical mapping and parametrized by a single competition parameter β. The dynamics of β is estimated here. *(ii)* We assume that the spike count in increasing time windows after stimulus onset forms the neural code. *(iii)* For each noise level $n(\vec{u}(t),t)$, $t \in [0, \Delta T]$ we estimate the optimal competition $\beta_t$ as if the input-output mapping was static in the $[0,t]$ interval. $\beta_t$ gives only an approximation for the optimal competition dynamics $\beta(t)$. An additive Gaussian noise on the estimated firing rate and a Poisson firing process are considered. For the Gaussian noise the variance of the estimated firing rate decreases with time. For the Poisson

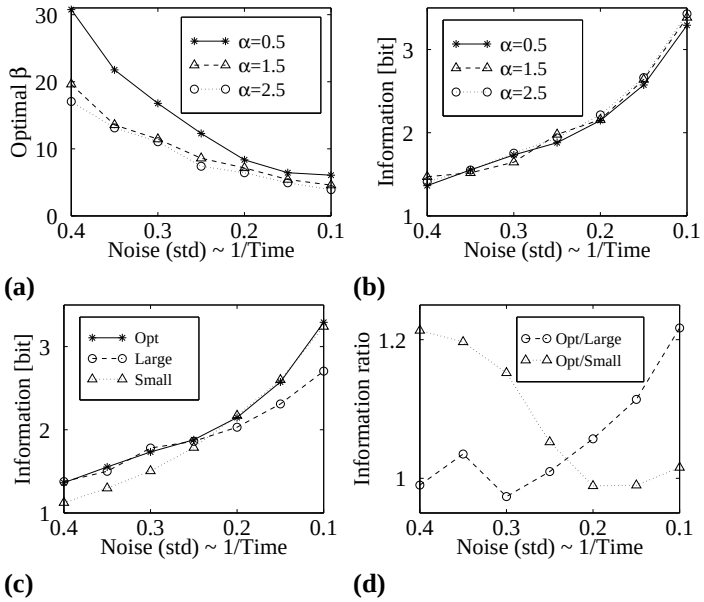


Figure 4.6: Information transfer if the output noise is additive Gaussian. **(a)** The optimal competition parameter $\beta$ as a function of the standard deviation of the additive Gaussian noise for different stimulus distributions. Small $\alpha$ parameter means sparse (super-Gaussian) input prior. The x-axis can be identified also with time because of the inverse relationship between the standard deviation of the spike count and the length of the time period since the stimulus onset. **(b)** The information transfer as a function of the standard deviation of the additive Gaussian noise. This could be also interpreted as the cumulative mutual information as a function of time after stimulus onset. The information transfer is independent of the sparsity parameter $\alpha$ because the signal distributions have the same variance. **(c)** The information transfer as a function of the standard deviation of the Gaussian noise for optimal dynamic (solid line), static small (dotted line) and static large (dashed line) competition. The sparsity parameter is $\alpha = 0.5$. If the competition is static high, then after an early fast rise, the information transfer saturates because the network is not able to represent multiple features. If the competition is static low, the cumulative mutual information reaches its maximal value only slowly. The ratios between the optimal and static setups **(d)** reveal that the dynamic modulation of the competition results in a maximum 20% increase in the information transfer. Note that these results for Gaussian noise are very similar to results achieved for Poisson spiking Fig. 4.7.

spiking model, the mean spike count in the entire network increases with time. In both cases the signal-to-noise ratio increases.

The estimation procedure of the optimal competition is described in detail in appendix B. Here we describe the course of our learning algorithm for the Poisson spiking case. Figure 4.5 shows the stochastic estimation of the optimal $\beta$ for a sparse a priori stimulus distribution ($\alpha = 0.5$) and for a time interval where the mean spike count $\mu$ is 7. The two-dimensional projections of the input- and output-distributions are shown in the top row of the figure. Each dot in the input

distribution represents an input pattern that is represented as intensity values along the different feature dimensions. In the context of early visual processing the input can be interpreted as the feed-forward representation of the part of an image, that falls into the cortical neurons' receptive fields. The spike count of the output neurons (different feature detectors) is also represented by a dot that falls into a little square that is centered at the observed spike count. (Note that the dots are jittered around each spike count coordinate only for demonstration reasons.)

In the bottom row, the first graph shows the mutual information ($I[X,Y] = H[Y] - H[Y|X]$) as a function of the learning step. With the first steps the mutual information grows fast and then it saturates early. The estimation of $\beta$ is stopped if the mutual information is converged. Note that the $\beta$ does not saturate as a function of the learning step because we increase the learning rate on $\beta$ exponentially. This speeds up the convergence of the mutual information. The incorporation of this improvement into our learning algorithm was motivated by our preliminary calculations that showed long flat plateau of mutual information as a function of $\beta$. The output entropy $H[Y]$ and the conditional uncertainty of the output $H[Y|X]$ are plotted as a function of learning step in the middle graph in the row. Plotting the two components of the mutual information separately shows that with increasing competition (increasing $\beta$) the conditional uncertainty decreases faster then the output entropy. This explains why does the mutual information increase with recurrent competition in the case of Poisson spiking.

### 4.3.3 Information transfer in time—Results

Figure 4.6a shows the optimal competition parameter $\beta$ as a function of the standard deviation of the additive Gaussian noise for different input distributions. The input distribution is parameterized by the sparsity parameter $\alpha$ (Eq. 4.6). The standard deviation is inversely related to time, therefore the x-axis can be also interpreted as the time axis. The calculations show that high additive noise requires strong competition between the feature detectors, while low competition is optimal if small noise is present. In the early phase of processing the noise on the neural code is high because only a few spikes are available for representation. Our calculations show that in this initial phase high recurrent competition is optimal with respect of information transfer. In the later phase of processing more spikes are available for the representation if the readout mechanism can integrate its input sufficiently long. It follows that the signal-to-noise ratio increases. Our calculations show that in this case, i.e. in the later phase of processing, low recurrent competition is optimal. If the recurrent competition is modulated dynamically, then mutual information is increased by 20% compared to what can be archived by static low or static high competition (Fig. 4.6c,d).

The calculations also revealed that at a sparser stimulus distribution higher competition is better. Intuitively, sparse distribution means that it is unlikely that two features are simultaneously present. Therefore the optimal representation strategy is to assign all the resources to a single neuron which represents the most salient feature. If the noise is additive or Poisson, this strategy increases the SNR of the neuron's response. Assigning all the resources (all spikes that are available due to metabolic or neurophysiological constraints) to very few neurons results in a highly transient response (cf. Figs. 4.3, 4.4). In the view of our information theoretical results one can interpret this early transient rise of activity as a mechanism to "counteract" the low SNR that is a consequence of the short available time for the representation. In the later phase of neural representation more time and therefore more resources are available for representation. In this later phase the activity can be distributed among the different feature detectors. Competition turns to be cooperation.

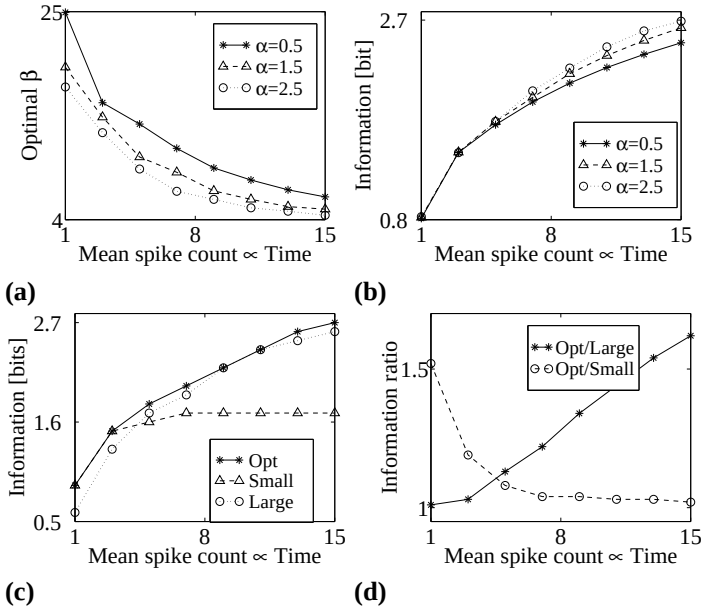


Figure 4.7: Information transfer if the output noise is multiplicative with a Poisson spiking statistics. **(a)** The optimal competition parameter $\beta$ as a function of the expected spike count for different stimulus distributions. Small $\alpha$ parameter means sparse (super-Gaussian) input prior. The x-axis can be identified also with time because assuming a constant overall network firing rate, the expected spike count is proportional to time. **(b)** The information transfer as a function of the expected spike count of the Poisson spiking units. This could be also interpreted as the cumulative mutual information as a function of time after stimulus onset. Note that our constraint on the variance of the signal distribution cannot ensure entirely that the mutual information remains independent of the sparsity parameter $\alpha$. This is different here compared to our results for the case of Gaussian noise. **(c)** The information transfer as a function of the expected spike count of the Poisson spiking units for optimal dynamic (solid line), static small (dotted line) and static large (dashed line) competition. The sparsity parameter $\alpha = 0.5$. If the competition is static high, then after an early fast rise, the information transfer saturates because the network is not able to represent multiple features. If the competition is static low, the cumulative mutual information reaches its maximal value slowly. The ratios between the optimal and static setups **(d)** reveal that the dynamic modulation of the competition results in a maximal 50% increase in the information transfer. Note that these results for Poisson spiking units are very similar to results achieved for additive Gaussian noise Fig. 4.6.

Similar results are obtained for the more realistic Poisson spiking model (Fig. 4.7). The optimal recurrent competition decreases with the expected spike count or equivalently with time. Our results are in accordance with previous theoretical results obtained for the limit cases of the expected spike counts (Brunel and Nadal, 1997). The dynamically modulated competition leads to a 50% improvement if compared to static low, or high competition (Fig. 4.7d, dotted and dashed lines respectively).

In the Poisson spiking model output entropy and the conditional uncertainty are monotonic decreasing functions of the recurrent competition parameter $\beta$ (cf. Fig. 4.5). The mutual information increases because the conditional uncertainty decreases faster with increasing competition than the output entropy does. This is different for the Gaussian noise model, where the output entropy as a function of recurrent competition has a maximum at finite $\beta$ and the conditional uncertainty is independent of the competition because of the additive nature of the noise. An intuitive explanation can be given for the different behaviors of the output entropies in the two noise models. For the Poisson spiking process high competition leads to an output distribution that is constrained onto the different axes that span the output space. In other words the distribution becomes effectively one-dimensional and therefore its volume and its entropy decreases towards zero. (Note that at high competition all neurons except one are inactive and their spike count has a zero variance assuming Poisson firing statistics.) The collapse of the output distribution to lines is a consequence of the multiplicative nature of the Poisson noise and it is not true if the noise is additive Gaussian. For the Gaussian noise model, increased competition distributes the centers of the Gaussian blobs that represent the mean firing rate of the single units. This increases the output entropy.

This theoretical study has important implications about the response dynamics and the dynamics of the SNR at visual cortical neurons. These predictions motivated by information theory are very similar to what have been obtained with the detailed computational study presented in section 4.2. Unfortunately there are hardly any experimental studies available which explore neural response dynamics and the time course of SNR in a free-viewing scenario or with flashed stimulus. Our results should be considered as model predictions that call for further experimental tests. For both noise models high competition in the initial phase of response leads to sharp onset transient. The neuron that represents the salient feature in the receptive field is very active, while other units are silent. In the later phase of the response, the activity is more distributed among units and may become lower to keep the energy consumption of the network constant over time. The transient increase in the early phase of cortical response is explained here by the objective to maximize information transfer. In the early phase, when only a few spikes can represent the input signal, high competition among the units maximizes the SNR. As the time window for representation increases, the activity becomes more distributed among the units: in the larger time window more spikes are available, therefore the network is capable to efficiently represent more complex input patterns. As a consequence of this strategy, the SNR—calculated for *fixed length* short sliding window—*decreases* with time after stimulus onset.[3] This is in full accordance with recent experimental observations obtained for primate visual cortical neurons (Heller et al., 1995; Wiener et al., 1999).

## 4.4 Discussion

In the present contribution we studied a yet unexplored aspect of short-term dynamics of synaptic strength and neural activity. We proposed that the neural processing strategy adapts to the internal variations of the nervous system itself: *dynamic neural code* emerges. We focused on the adaptation of visual cortical neurons to the non-stationary output noise on the time scale of a fixation period. Note that dynamic modulation of the neural code can be motivated by more complex, cognitive function related internal changes in the nervous system, for instance by modulations in attentive states.

[3] Note the difference between the SNR calculated for fixed length time windows and increasing time windows beginning with the stimulus onset.

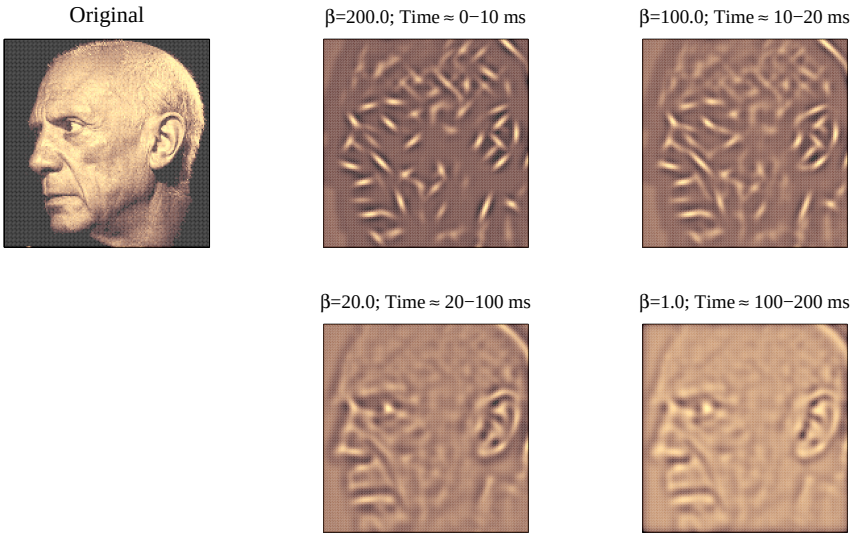


Figure 4.8: This is how we imagine neural representation of and image in V1. This schematic illustration shows four snapshots of the cortical representation at different moments within a fixation period. The original image is filtered with elongated, simple cell like receptive field profiles with different elongation orientations. The filtered image at a certain orientation represents the feed-forward input to an orientation selective simple cell in V1. The input to units with all orientations at neighboring positions is mapped by a soft-max function with a time dependent competition parameter $\beta$ that is indicated at each image. The result can be interpreted as the neural output. The neural activity is then back-projected to the neurons' receptive field profiles. The result of this back-projection is shown in the figures.

### 4.4.1 Model conclusions

Making the assumption that the entire neural activity pattern after stimulus onset carries the information about the stimulus, we suggested that the SNR of the neural code decreases on a short time scale. In the initial period of response after stimulus onset, only a few spikes are available for encoding because the dynamic range of the neural activity is limited. As the time window for representation becomes wider, more and more spikes can "describe" the input in a more precise manner. We have demonstrated that given this time course of the noise on the neural representation, maximal information transfer in any time window after stimulus onset can be achieved if the competition between feature detectors is initially high and then decreases with time. Intuitively this means that the network first extracts the salient and typical features, which carry the most information on average about the input. Details are processed in the second phase. Obtaining maximal information about the stimulus in any period could be essential for survival because it allows for fast reaction times. Therefore it is not unlikely that it is developed evolutionary. Similar processing strategy in primate temporal visual cortex has been demonstrated experimentally recently (Sugase et al., 1999).

The general question of maximizing information transfer in time if the noise on the output units is non-stationary was studied here in the context of orientation selectivity in the primary visual

cortex. We have set up a computational model for a cortical hypercolumn. In the model simple cells are connected recurrently with orientation specific, fast adapting excitatory and non-specific, static inhibitory connections. Due to the fast synaptic depression at the lateral excitatory connections the initially strong recurrent competition decreases. In our recent study we propose spike frequency adaptation of regular spiking neurons as an alternative mechanism that could implement this dynamic coding strategy (Schwabe et al., 2000). Our theoretical results show, that decreasing the recurrent cortical competition after stimulus onset follows naturally from the requirement that the mutual information between stimuli and representations is maximal for any time interval beginning with stimulus onset. Furthermore, the dynamic recurrent network obtains both contrast invariant sharpening of the feed-forward input and the faithful representation of multiple features. The dynamic recurrent amplifier model forms a compromise between the "feed-forward" and the "recurrent" hypotheses for orientation selectivity and may resolve contradictory experimental results regarding the role of recurrent cortical processing in the generation of orientation selective response. Sharp contrast invariant orientation tuning is generated in the initial phase of cortical response when the recurrent connections are strong enough. In that phase the salient feature is signaled in a winner-takes-all fashion. Finer details are signaled reliably in the second phase cortical response.

The predicted time course of the visual cortical representation of an image is depicted in Fig. 4.8. These schematic calculations demonstrate that attenuation of the competition between edge detectors obtains a hierarchical extraction of features on different levels of complexity. The hierarchical levels are distributed in time. First the salient features, the "keywords" of the image are extracted. Details are processed only in the later phase. Similar strategy could be employed in artificial communication systems where a signal has to be transmitted online through a channel with limited capacity.

### 4.4.2 Model assumptions

The key assumption in the current study is that the entire activity pattern within a period that is longer than the average inter-spike-interval constitutes the neural code about a stimulus. To justify this assumption, new experiments have to be designed to explore the read-out mechanisms in the cortical information channels. Given the constraints on the maximal firing rate and the imprecision of spike timing, here we proposed a relatively complex dynamic code. The exploration of possible read-out processes was out of the scope of the current study. On the one hand, the here derived dynamic code is optimal in the sense of information transfer and therefore potentially allows for the most efficient readout. On the other hand, neurophysiological constraints at the read-out mechanism could make suboptimal encoding easier to access.

Another important aspect of our investigation is that we considered a "free-viewing" scenario with "saccading eye movements" modeled as a series of visual stimuli that are flashed onto the retina for the duration of a "fixation period"($200 - 300$ ms). We made the assumption that the stimuli and the feed-forward input to the primary visual cortex are static in order to be able to present a clear-cut model for the short term cortical dynamics that is only due to the time course of the output noise. Inclusion of dynamic stimuli and the phasic transient input that arrives from the geniculate X-cells could be a focus of some future studies.

We showed that optimal information transfer is obtained if the recurrent competition is strong in the beginning of a fixation period and then it is attenuated. Fast synaptic depression due to the depletion of synaptic resources could be a neurophysiological mechanism to implement this coding strategy. In other words, strong recurrent competition arises because initially larger amount of

synaptic transmitter is available. Towards the end of a fixation period the synaptic resources are used up and therefore the recurrent competition attenuates. It follows however, that the synapses of previously active neurons have to recover very fast during a saccade to ensure strong competition in the beginning of the next fixation period. However, this is unrealistic because the time constant for transmitter recovery is in the range of a few hundred milliseconds (Abbott et al., 1997; Tsodyks and Markram, 1997). This paradox can be easily resolved if we consider that the edges with different orientations form a sparse distributed representation of natural scenes (Field, 1994; Ruderman, 1994). This means that it is relatively unlikely that neurons receive strong feed-forward input in subsequent fixation periods and neurons that have been previously silent can obtain an efficient strong competition. Furthermore, experimental studies reveal that the saccade and the flashed stimulus elicited neural responses in V1 are very different (Leopold and Logothetis, 1998). This observation allows to speculate that there are "reset" mechanisms present that are controlled over motor feed-back initiated by a saccading eye-movement.

### 4.4.3 Model predictions

The model has important predictions about the fast neuronal activity dynamics in V1 and about the dynamics of perception of flashed static stimuli. Our simulations suggest that the visual cortical signaling of a low intensity edge is delayed if it is masked by a second higher intensity edge. Further simulations (Schwabe, 1999) showed that the delay scales with the intensity difference between the two edges. Averaging over the entire activity course within a fixation period, we also found that in accordance with the experimental results (Hata et al., 1988; Bonds, 1989; DeAngelis et al., 1992) masking suppresses the activity strength. To sum up, our model predicts that a masking stimulus modulates both the strength and the dynamics of the cortical response. This prediction could be directly tested by measuring single cell activity extracellularly in the primary visual cortex for flashed bar and cross shaped stimulus. Psychophysical investigations could also test for our model hypothesis. Our model predicts that if a mixture of static features is flashed for a short interval, then only the salient is perceived even though the weaker feature presented alone is perceived correctly within that period. Our results also suggest that long-range integration arriving from outside the "classical receptive" field happens in the second phase of processing similar to the signalling if weak intensity details within the classical receptive field.

The model has important predictions with respect to the fast dynamics of information transfer in visual cortical neurons. Mutual information is the most general qualitative measure of the "quality" of representation. However, it is hard to estimate if only a limited number measurements are available. An approximative view about information transfer can be gained, e.g., by considering only the signal-to-noise ratio. We suggest that a transient rise of activity emerges in the initial phase of cortical processing because of the high competition between cortical cells. As a consequence the mutual information rises fast and the SNR is high. In the second phase, the competition between feature detectors decreases, the network represents the unusual and small details if they are present. Representing complex mixtures of different features requires the distribution of activity. This leads to a decrease in the SNR if it is measured for the activity within a fixed width sliding time window. This is in accordance with a current experimental observation (Wiener et al., 1999). Note that the SNR and the cumulative mutual information between input and output increases if the cortical activity for a time window with increasing length is taken into account (see Figs. 4.6, 4.7). The simulations show that a fast rise is followed by a slow saturation that has been also observed experimentally (Heller et al., 1995).

# Chapter 5

# Contrast adaptation and infomax in visual cortical neurons


## Abstract

In the primary visual cortex (V1) the contrast response function of many neurons saturates at high contrast and adapts depending on the visual stimulus. We propose that both effects—contrast saturation and adaptation—can be explained by a fast and a slow component in the synaptic dynamics. In our model the saturation is an effect of fast synaptic depression with a recovery time constant of about 200 msec. Fast synaptic depression leads to a contrast response function with a high gain for only a limited range of contrast values. Furthermore, we propose that slow adaptation of the transmitter release probability at the geniculo-cortical synapses is the underlying neural mechanism that accounts for contrast adaptation on a time scale of about 7 sec. For the functional role of contrast adaptation we make the hypothesis that it serves to achieve the best visual cortical representation of the geniculate input. This representation should maximize the mutual information between the cortical activity and the geniculo-cortical input by increasing the release probability in a low contrast environment. We derive an adaptation rule for the transmitter release probability based on this *infomax* principle. We show that changes in the transmitter release probability may compensate for changes in the variance of the geniculate inputs—an essential requirement for contrast adaptation. Also, we suggest that increasing the release probability in a low contrast environment is beneficial for signal extraction, because neurons remain sensitive only to an increase in the presynaptic activity if it is synchronous and, therefore, likely to be stimulus related. Our hypotheses are tested in numerical simulations of a network of integrate-and-fire neurons for one column of V1 using fast synaptic depression and slow synaptic adaptation. The simulations show that changing the synaptic release probability of the geniculo-cortical synapses is a better model for contrast adaptation than the adaptation of the synaptic weights: only in the case of changing the transmitter release probability our model reproduces the experimental finding that the average membrane potential (DC component) adapts much stronger than the stimulus modulated component (F1 component). In the case of changing synaptic weights, however, the average membrane potential (DC) as well as the stimulus modulated component (F1 component) would adapt. Furthermore, changing the release probability at the recurrent cortical synapses cannot account for contrast adaptation, but could be responsible

for establishing oscillatory activity often observed in recordings from visual cortical cells[1].

## 5.1 Introduction

The contrast response function of simple cells in the primary visual cortex (V1) saturates at high stimulus contrast and neurons show a high gain at a limited range of contrast only. The contrast response function of most neurons in V1 adapts to slow changes in the visual environment by shifting the part of the contrast response function with the highest gain towards the most typical contrast level presented in the preceding few seconds. The contrast response function shifts towards higher contrast values following a prolonged presentation of high contrast preferred stimuli. It shifts to the opposite direction if the preceding stimuli had low contrast for several seconds. This adaptation makes it possible to efficiently map the widely changing input signals to an output with a limited dynamic range. Contrast adaptation is essentially a cortical phenomenon. Lateral geniculate cells that provide the main feed-forward input to V1 partially (Shou et al., 1996) or do not adapt to contrast in cat (Ohzawa et al., 1985; Bonds, 1991; Ahmed et al., 1997) and monkey (Sclar et al., 1989).

During the last decade several ideas emerged to explain contrast adaptation, but none of them was fully consistent with the experimental data. On the one hand, a group of studies suggests that plasticity of excitatory synaptic efficacy or weights[2] is responsible for contrast adaptation (Barlow and Földiák, 1989; Todorov, Siapas, Somers and Nelson, 1997; Chance et al., 1998). On the other hand, the dynamical properties of the cellular response change after contrast adaptation and modulation of synaptic *weights* cannot account for these effects. To resolve the abovementioned contradictions, we propose the new hypothesis that contrast adaptation is the result of changing the dynamic properties of the synapses. In particular, we suggest that the slow modulation of the transmitter release probability plays a key role. Furthermore, we suggest that the transmitter release probability changes in order to maximize the mutual information between the geniculate input and the cortical output, that is, according to the *infomax* principle.

Let us first review the experimental results in detail. One possible mechanism for contrast adaptation is the intracellular adaptation of neurons. However, fatigue of cortical cells is unlikely to be the origin of contrast adaptation, because adaptation is not effected by altering the firing rate of a neuron (DeBruyn and Bonds, 1986; Vidyasagar, 1990). Recurrent inhibition is not likely to be a key factor in contrast adaptation either, because no significant alterations in membrane conductance are found after contrast adaptation or with increasing stimulus contrast (Carandini and Ferster, 1997)[3]. This indicates that shunting (divisive) inhibition does not play an important role in cortical gain control and thus in contrast adaptation. Carandini and Ferster (1997) concluded that a decrease in tonic excitation rather than an increase in inhibition accounts for contrast adaptation. This view concurs with other experimental results showing that local inactivation of GABAergic inhibitory synapses has no effect on contrast adaptation (DeBruyn and Bonds, 1986; Vidyasagar, 1990; McLean and Palmer, 1996), whereas blocking glutamate (excitatory) autoreceptors that mediate synaptic depression decreases the degree of adaptation (McLean and Palmer, 1996). Furthermore, adaptation

[1] This chapter is based on (Stetter et al., 1998; Adorján and Obermayer, 1999; Adorján, Piepenbrock and Obermayer, 1999)

[2] Throughout this paper by the term "synaptic weight" or "synaptic strength" we mean every synaptic parameter that scales *only* the amplitude of the postsynaptic current. This could be, the quantal size, the absolute amount of released vesicles or the maximal postsynaptic conductance. We distinguish these physiological parameters from those that modify the synaptic behavior in a more complex manner like the transmitter release probability.

[3] The issue of lateral inhibition is still controversial. Recent data (Borg-Graham et al., 1998) show large stimulus evoked changes in the input conductance and suggest that conductance changes may have been underestimated in the past.

depends on the stimulus pattern: it is strongest if the adapting and the testing stimuli are the same (Movshon and Lennie, 1979; Saul and Cynader, 1989*a*; Saul and Cynader, 1989*b*; Carandini, Barlow, O'Keefe, Poirson and Movshon, 1997; Carandini et al., 1998). These observations strongly indicate that plasticity of excitatory synapses, rather than the inhibitory synapses, plays a key role in contrast adaptation.

Following this idea, several modeling studies proposed adaptation of the excitatory synaptic *weights* for contrast adaptation. Recently, short term synaptic depression was suggested (Todorov, Siapas, Somers and Nelson, 1997), although the time constant for synaptic depression (200-800 msec) (Tsodyks and Markram, 1997; Abbott et al., 1997) is one magnitude smaller than the time constant for contrast adaptation (5-10 s) (Ohzawa et al., 1985). However, changing the synaptic efficacy on a longer time scale (Chance et al., 1998) cannot account fully for contrast adaptation either. Surprisingly, the dynamics of the neural response change with contrast adaptation. Two puzzling phenomena have been reported. First, the response delay increases after high contrast adaptation (Saul, 1995). Second, the average membrane potential (DC component) adapts stronger to contrast than the stimulus driven modulation of the membrane potential (F1 component) (Carandini and Ferster, 1997). Adaptation by changing the synaptic *weights* (Todorov, Siapas, Somers and Nelson, 1997; Chance et al., 1998) cannot explain these effects.

In summary, $(i)$ contrast adaptation is more likely to be due to a decrease in tonic excitation than an increase in recurrent inhibition and $(ii)$ contrast adaptation seems to involve activity dependent synaptic plasticity that $(iii)$ changes the dynamical properties of synaptic transmission rather than the synaptic strength. The dynamic properties of the synaptic transmission depend strongly on the transmitter release probability and therefore it could be the synaptic property that is modulated during contrast adaptation. Furthermore, a series of in vitro experiments demonstrated that presynaptic modulation of the transmitter release is a very likely mechanism for adaptation in cortical cells (Finlayson and Cynader, 1995). These considerations motivated us to explore the hypothesis that changing the transmitter release probability of the excitatory synapses accounts for contrast adaptation.

The second hypothesis of this paper is that the response of cortical neurons should maximize the information throughput from the retina to the primary visual cortex. This assumption is motivated by the experimentally observed nature of contrast adaptation. On the time scale of $\tau \approx 7\,\mathrm{sec}$ cortical neurons adapt to the slowly changing average contrast of the visual environment by shifting the sensitive regime of their transfer function to the input values with the highest probability density. Maximizing the mutual information (Shannon and Weaver, 1949) between input and output makes optimal use of the *limited* encoding bandwidth in a similar way. Contrast adaptation eliminates the dependence of the cortical activity on the *global* contrast level that soon after a change represents redundant information. This increases the neuronal sensitivity to the *local* contrast fluctuations with the most likely amplitude or, in other words, to the image pattern that carries the important information. Light adaptation is likely to serve a similar goal by making the activity of the retinal ganglion cells independent of the mean luminance (Shapley and Enroth-Cugell, 1984). The early visual system adapts to the first (mean luminance) and to the second order (global contrast) statistics of the actual visual environment in a two stage process keeping the information mostly about the higher order statistics (image pattern).

We derive the adaptation rule for the transmitter release probability from the infomax principle: it maximizes the information content of the cortical activity about its input and therefore reduces the redundancy in the neural activity. Note again that this is a sensible goal only if the coding capacity of the cortical units is limited, that is, their output is noisy and its intensity range is constrained. Clearly,

an ideal noiseless neuron could encode an infinite amount of information independent of its transfer function. In contrast, a noisy neuron without adapting to the most typical contrast level would redundantly respond with similar, hardly distinguishable firing rates to most of the visual stimuli. Cortical representation is improved by matching the high gain part of the transfer function to the most likely input signals. This principle of matching the neuronal transfer function to the distribution of signals was first explored experimentally by Laughlin (1981). The principle of removing redundancy present in the natural visual signals in order to obtain an efficient neural representation in the early visual system was investigated by several studies before (Barlow, 1961; Barlow and Földiák, 1989; Linsker, 1989; Atick and Redlich, 1990).

In section 5.2 we set up a computational model for a small segment (one orientation column) of V1 with cortical simple cells represented by integrate-and-fire units with fast depressing synapses (Abbott et al., 1997; Tsodyks and Markram, 1997). In section 5.3.1 we derive an adaptation rule for the transmitter release probability of the geniculo-cortical synapses. The adaptation rule is based on an infomax principle to maximize the mutual information between the geniculate input and the cortical output activity. Furthermore in section 5.3.2, by discussing an analytical approximation for the synaptic current, we show that an increase in the transmitter release probability reduces the temporal window for input summation and makes a neuron sensitive only to a synchronized increase in the presynaptic firing rates. As synchronous changes in the geniculate activity are likely to be stimulus related, this is advantageous for extracting signals in a low contrast environment. This theoretical result also demonstrates that a change of the release probability is a better adaptation model than a change of the synaptic weight. In section 5.4 we present numerical simulations incorporating both fast synaptic depression and contrast adaptation by changing the transmitter release probability. The computational results are compared with the available experimental observations. A preliminary short report on our results has been published elsewhere (Adorján and Obermayer, 1999).

## 5.2 Model setup

### 5.2.1 The neural network

The model has one layer that corresponds to a cortical orientation-column with strong local recurrent connectivity (Fig. 5.1). In the model all 30 leaky-integrator excitatory units are recurrently connected with each other. Model neurons represent simple cells in the thalamic recipient zone of V1. The geniculate input to cortical units is described by time modulated independent Poisson spike trains arriving at each synapse. The time modulation of the firing rates reflects the changes of the visual stimulus at a single point in the visual field. In order to investigate the hypothesis that adaptation of the excitatory synapses alone can account for the saturation of the contrast response function and for contrast adaptation, recurrent inhibition is not included in the model. In the following we describe our model in detail.

### 5.2.2 Single cell and synaptic model

The model network consists of leaky integrate-and-fire neurons. Between spikes, a neuron's membrane potential increases due to its integrated excitatory synaptic input and decreases due to leakage

$$C_{\mathrm{m}}\frac{\partial V_i(t)}{\partial t} \quad = \quad I_{\mathrm{leak}} + I_{\mathrm{syn}} = g_{\mathrm{leak}}(E_{\mathrm{rest}} - V_i(t)) + \sum_j g_{ij}(t)(E_{\mathrm{syn}} - V_i(t)) \;, \tag{5.1}$$

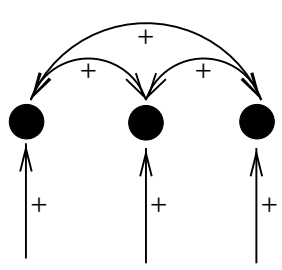


Figure 5.1: The model architecture.

where $C_{\mathrm{m}} = 0.5\,\mathrm{nF}$ is the membrane capacitance, $g_{\mathrm{leak}} = 31\,\mathrm{nS}$ is the membrane leakage conductance, $E_{\mathrm{rest}} = -65\,\mathrm{mV}$ is the membrane resting potential, and $E_{\mathrm{syn}} = -5\,\mathrm{mV}$ is the excitatory synaptic reversal potential. When the membrane potential reaches the firing threshold of $-55\,\mathrm{mV}$, the neuron emits a spike and the membrane potential $V_i(t)$ is reset to 1 mV below the resting potential after an absolute refractory period of 2 msec. The effect of a spike from neuron $j$ on target cell $i$ is modeled by the alpha-function: after a presynaptic spike at time $t_j^s$ the postsynaptic conductance increases rapidly to its maximal value of $g_{\max} \cdot p_{ij}(t_j^s) \cdot R_{ij}(t_j^s)$ and then decays slower

$$g_{ij}(t) \quad = \quad \frac{g_{\max}}{\tau_{\mathrm{peak}}} \sum_{s}^{|\mathrm{spikes}|} p_{ij}(t_j^s) \cdot R_{ij}(t_j^s) \cdot (t - t_j^s) \cdot \exp\left(1 - \frac{t - t_j^s}{\tau_{\mathrm{peak}}}\right) , \tag{5.2}$$

where $g_{\max} = 7.8\,\mathrm{nS}$ is the maximum conductance, $\tau_{\mathrm{peak}} = 1\,\mathrm{ms}$ is the rise-time of the alpha-function, $p_{ij}$ the transmitter release probability, $R_{ij}$ is the amount of available synaptic transmitter (synaptic "resource", Tsodyks and Markram, 1997), and $t_j^s$ is the arrival time of spike number $s$ from neuron $j$. It is important to note that the maximum amplitude of the synaptic conductance in this model is not a static property of a synapse, but it is proportional to the amount of synaptic transmitter $p_{ij}(t_j^s) \cdot R_{ij}(t_j^s)$ released after the spike. The above model parameters correspond to a membrane time constant of 16 msec (McCormick et al., 1985; Stratford et al., 1996) and an average EPSP amplitude of 1 mV (Stratford et al., 1996) measured in regular-spiking and spiny stellate excitatory neocortical neurons. Equations 5.1 and 5.2 constitute a basic integrate-and-fire model framework.

Fast synaptic depression is modeled by the dynamics of the synaptic transmitter or "resource" $R_{ij}(t)$. The maximal amount of synaptic transmitter is assumed to be finite at each synapse. As long as no presynaptic spike arrives, the amount of transmitter recovers exponentially to 1 with a time constant $\tau_{\mathrm{rec}} = 200\,\mathrm{ms}$. Upon each arrival of a presynaptic spike the resource is depleted proportionally to the transmitter release probability $p_{ij}(t)$ by an amount of $p_{ij}(t) \cdot R_{ij}(t)$. The depletion occurs on a millisecond time scale (much faster than $\tau_{\mathrm{rec}}$). The dynamics of the synaptic transmitter $R_{ij}(t)$ for the synapse between the presynaptic neuron $j$ and the target neuron $i$ follow the stochastic differential equation

$$R_{ij}(T + \Delta) - R_{ij}(T) \quad = \quad \int_T^{T+\Delta} \frac{1 - R_{ij}(t)}{\tau_{\mathrm{rec}}}\, dt - \int_T^{T+\Delta} p_{ij}(t) \cdot R_{ij}(t)\, dN(t) , \tag{5.3}$$

where $\tau_{\mathrm{rec}} = 200\,\mathrm{ms}$ (Tsodyks and Markram, 1997) is the recovery time constant and $N(t)$ is a spike counter that is incremented by one at the arrival of the presynaptic spikes. The solution of this

stochastic differential equation in the *Itô sense* between spikes satisfies the deterministic differential equation

$$\frac{dR_{ij}(t)}{dt} = \frac{1 - R_{ij}(t)}{\tau_{\text{rec}}}, \tag{5.4}$$

and in the neighborhood of a spike at time $t_s$ the solution is

$$\lim_{\substack{t \to t_s \\ t > t_s}} R_{ij}(t) = (1-p) \lim_{\substack{t \to t_s \\ t < t_s}} R_{ij}(t) . \tag{5.5}$$

This model of fast synaptic depression may be studied for slowly changing presynaptic firing rates (see equation C.3) and the steady state approximation for the amount of transmitter predicts that the synaptic transmission saturates for high presynaptic firing rates. Thus, it shows a high gain only for a limited range of presynaptic firing rates and this range depends on the transmitter release probability. This property is used in the derivation of the slow adaptation rule for the transmitter release probability $p_{ij}(t)$ that we propose for contrast adaptation in the next section.

## 5.3 Slow dynamics and contrast adaptation—theoretical results

### 5.3.1 Adaptation rule—infomax

We now assume that a cortical neuron adapts to a new image contrast by slowly changing the transmitter release probability at its feed-forward synapses. Here we follow our hypothesis that contrast adaptation pursues the objective to maximize the information that a cortical neuron's output conveys about its geniculate input. In other words, the goal is to maximize the *mutual information* $I[O, f]$ between a visual cortical neuron's output firing rate $O(f, p)$ and its geniculate input firing rate $f$. Maximizing the mutual information makes optimal use of the limited neuronal bandwidth by making each output value equally likely or, in other words, by maximizing the output entropy while minimizing the uncertainty of the output once the input signal is fixed. In the derivation we assume that the uncertainty in the neural mapping cannot be reduced and it is additive in nature.

A flat output probability density can be obtained by matching the neural transfer function to the input densities. This requires an estimation of the input distribution. Image contrast is defined as the standard deviation of the luminance values normalized by the mean luminance (RMS contrast, Shapley and Enroth-Cugell (1984)). Light adaptation in the retina makes the retinal input to the LGN largely invariant of the mean luminance. However, the information about the global image contrast is preserved. This statistical property of the visual environment is reflected mainly in the variance of the firing rates[4] of the geniculate cells with different receptive field positions. Now let us consider a sampling of the visual environment in time by fixating at different positions or saccading around, and again consider the firing rates of a geniculate cell for each of these fixation periods. In this scenario the global image contrast will be reflected mainly in the variance of the firing rates over the recent fixation periods. Thus, from a statistical point of view, for a cortical neuron contrast adaptation is "as easy" as estimating the variance of the presynaptic geniculate firing rates during the recent fixation periods. A single cortical neuron has to adapt to the image contrast by observing typical contrast edges in the whole environment. When looking around freely, the eyes saccade 3 to 4

[4] The firing rate for one fixation period.

times a second (Viviani, 1990; Stetter et al., 1996). This means that a cortical neuron gets the chance to sample the local contrast within its receptive field at about 25 image locations within 7 sec—a sufficient number to coarsely adapt to the global image contrast. For the upcoming simulations we consider the experimental conditions and use grating stimuli with fixed eye position. But note that the adaptation rule we derive below solves the more general problem of adapting the neural transfer function to match the distribution of any type of input signals. Thus it is a proper mechanism for pattern adaptation.

We derive an adaptation rule for the release probability $p$ given the output firing rate $O(f,p)$ of a neuron in the recurrent network (estimated in equation C.4) as a function of the presynaptic firing rate $f$ and the release probability $p$. This adaptation rule is then incorporated into the integrate-and-fire neural network model and studied by numerical simulations in section 5.4. For fast changing presynaptic firing rates, we investigate the nature of synaptic transmission in the following section.

The mutual information (Shannon and Weaver, 1949) is defined as the difference between the entropy $H[O]$ of the output and the conditional entropy $H[O|f]$ of the output given the input:

$$I[O(f,p),f] \quad = \quad H[O(f,p)] - H[O(f,p)|f]\,. \tag{5.6}$$

This means that a coding is efficient if it has a high variability if all possible inputs are taken into account, but shows high fidelity once the input is fixed. Maximizing equation 5.6 is equivalent to maximizing the entropy of the output if the input-output mapping $O(f,p)$ is deterministic or if we assume only additive noise (Bell and Sejnowski, 1995). Thus our objective is to find the transmitter release probability $p$ that maximizes

$$\begin{aligned} H[O(f,p)] &= -E\Big[\ln \mathrm{Prob}(O(f,p))\Big] &(5.7)\\ &= -E\Big[\ln \frac{\mathrm{Prob}(f)}{|\partial O(f,p)/\partial f|}\Big] &(5.8)\\ &= E\Big[\ln\Big|\frac{\partial O(f,p)}{\partial f}\Big|\Big] - E\big[\ln \mathrm{Prob}(f)\big]\,, &(5.9)\end{aligned}$$

where $E[.]$ denotes the expectation value over all input rates $f$ given some image contrast and $\mathrm{Prob}(.)$ is the probability density function of its argument. The second term in equation 5.9, the entropy of the presynaptic firing rate $f$ is independent of the parameter $p$. Therefore the entropy of the cortical output can be maximized by gradient ascent only on the first term of equation 5.9. Note that this entropy depends on the probability density of the presynaptic firing rate $\mathrm{Prob}(f)$ that has to be estimated in an online manner. An online adaptation rule is obtained by considering random samples from the presynaptic firing rate distribution (Bell and Sejnowski, 1995). With the approximation for $O(f,p)$ from equation C.4 this yields a stochastic adaptation rule for the transmitter release probability

$$\begin{aligned} \tau_{\mathrm{adapt}}\frac{\partial p}{\partial t} &= \frac{\partial}{\partial p}\ln\Big|\frac{\partial O(f,p)}{\partial f}\Big| \\ &= -2\tau_{\mathrm{rec}} fR + \frac{1}{p} + \frac{\tau_{\mathrm{rec}}(fa-1)}{a+\tau_{\mathrm{rec}}p(fa-1)}\,, &(5.10)\end{aligned}$$

where the time constant $\tau_{\mathrm{adapt}} = 7\,\mathrm{sec}$ is adjusted to the measured time constant of contrast adaptation (Ohzawa et al., 1985) and $a = \frac{\alpha+1}{f} - \frac{1}{f+\Theta}$ with parameters $\alpha$, $\Theta$ that are related to the description

of the nonlinear summation of subsequent EPSC pulses (see appendix C.2). The first term in the adaptation rule is proportional to the presynaptic firing rate $f$ and to the available synaptic transmitter $R$ suggesting a local presynaptically driven non-Hebbian mechanism for adaptation. The amount of synaptic transmitter $R$ determines the amplitude of the EPSC that potentially could be evoked by a presynaptic spike. The probability of the transmitter release $p$ decreases if the presynaptic firing rate and the EPSC that potentially could be evoked are high. The second term ensures that $p$ is always larger than 0 and in this model setup $p$ also stays always below 1. The third term modulates the adaptation slightly and increases the release probability $p$ most if the input firing rate is close to 20 Hz, that is at low image contrast.

### 5.3.2 Redistribution of synaptic resources—transients

The cortical neurons show transient responses following fast changes in the presynaptic firing rates due to the fast depressing synapses. In the previous section we have derived an adaptation rule for slowly changing inputs. In this section we analyze the contribution of the transients to the total cortical response by investigating the dynamics of the excitatory postsynaptic current (EPSC). We will show that changes in the transmitter release probability may compensate for changes in the variance of the geniculate inputs—an essential requirement to make contrast adaptation work. Also, we suggest that increasing the release probability in a low contrast environment is beneficial for signal extraction because neurons remain sensitive only to an increase in the presynaptic activity if it is synchronous and, therefore, likely to be stimulus related. The results presented in this section could constitute the basis of a more realistic transfer function and an adaptation rule that takes the transient synaptic current also into account.

To study the effect of transients, we consider a free viewing scenario with fixation periods followed by saccades. For the analysis within this framework we use the following assumptions. $(i)$ A cortical synapse receives an input firing rate $f(t)$ that stays largely constant during a fixation period and changes suddenly at each saccade. $(ii)$ The fixation periods between two saccades (200 msec to 400 msec (Viviani, 1990; Stetter et al., 1996)) are long enough for the amount of synaptic transmitter to reach its steady state $R_f^\infty$. $(iii)$ The duration of a saccade is short compared to the duration of a fixation period and we neglect it. In summary, the presynaptic firing rate changes as a step function with saccades and during a fixation period only the amount of available synaptic transmitter changes. Other factors—the maximal conductance $g_{\max}$ and the transmitter release probability $p$—that determine the EPSC stay constant on the time scale of a fixation period. It follows, that to investigate the dynamics of the EPSC, we need to focus on the dynamics of the available amount of transmitter.

The basis of our investigation is the differential equation that describes the dynamics of the *expected* amount of transmitter $\bar{R}(t)$—equation C.1, where the expectation is calculated over different realizations of a Poisson-process firing with a given rate. We solve equation C.1 for a fixation period, during which the presynaptic firing rate $f^{\text{new}}$ is assumed to be fixed (first assumption for the saccadic framework). Noting that a fixation period is long enough and the saccade itself is short (second and third assumptions), it follows that the previous steady state $R_{\text{old}}^\infty$ can serve as the new initial condition $\bar{R}(t_0)$ and the resource dynamics is a function of the two consecutive afferent firing rates $f^{\text{old}}$ during the previous and $f^{\text{new}}$ during the present fixation period:

$$\bar{R}(f^{\text{old}}, f^{\text{new}}, t) \quad = \quad \left(R_{\text{old}}^\infty - R_{\text{new}}^\infty\right) \exp\left(-\frac{t-t_0}{\tau_{\text{rec}} R_{\text{new}}^\infty}\right) + R_{\text{new}}^\infty \,. \tag{5.11}$$

This expression has two terms: a transient component that decays with the effective time constant

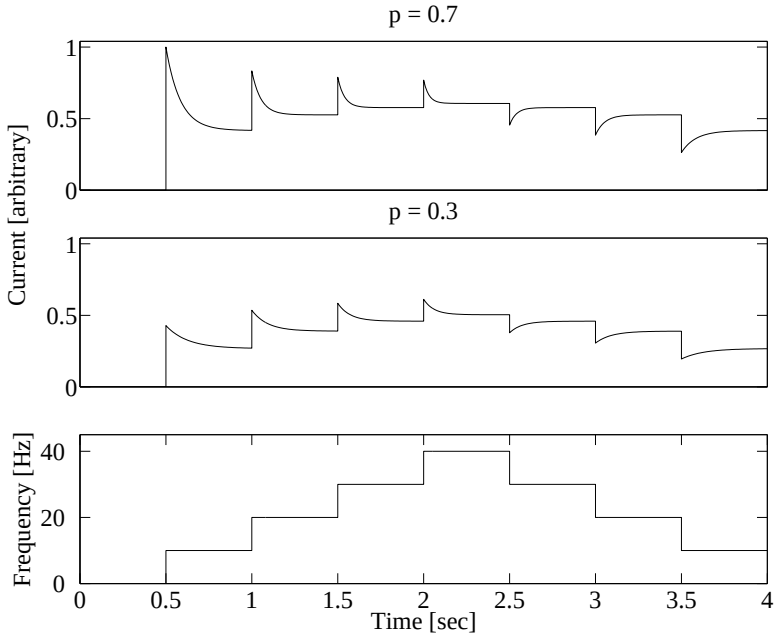


Figure 5.2: The postsynaptic current (equation 5.12, normalized) as a function of time at high (top panel) and low (middle panel) transmitter release probability. The presynaptic firing rate is plotted on the bottom panel.

of $\tau_{\text{eff}}(f^{\text{new}}, p) = \tau_{\text{rec}} R^{\infty}_{\text{new}}$ and a steady state $R^{\infty}_{\text{new}} = 1/(1 + p f^{\text{new}} \tau_{\text{rec}})$. For the EPSC we obtain the approximation

$$I(f^{\text{old}}, f^{\text{new}}, t) \propto g_{\max} f^{\text{new}} p \bar{R}(f^{\text{old}}, f^{\text{new}}, t) \; . \tag{5.12}$$

The time evolution of the synaptic current for two different transmitter release probabilities is shown in figure 5.2. The plot demonstrates that if the presynaptic firing rate changes abruptly, the transient part of the EPSC dominates the synaptic input. Therefore the transient current determines the neuron's activity to a large extent. The input-output relationship of a neuron can be well approximated by considering the mean current in the first short transient phase following the jump or the saccade. Note that this current depends on the presynaptic firing rate both before and after the saccade.

The steady state component of the current saturates at high presynaptic firing rates. The larger the release probability is, the more the amplitude of the transient current scales with the *relative* jump ($(f^{\text{new}} - f^{\text{old}})/f^{\text{old}}$) in the presynaptic firing. With increasing release probability the amplitude of the transient current and the average current grows. At low release probability the synapse is less sensitive to input transients, while at high release probability an increase in the input activity rate results in large and sharp transient EPSC. Consequently, in a high contrast environment when the presynaptic geniculate firing rate makes large jumps between fixation periods, a cell receives approximately the same transient EPSC as at low contrast when the jumps are smaller in the presynaptic firing rate. It follows that changes in the transmitter release probability that compensate for variances in the input rates could serve as a neural mechanism for contrast adaptation.

In contrast to the overall synaptic strength $g_{\max}$, the release probability does not simply scale the amplitude of the synaptic input. The release probability also determines the effective time constant $\tau_{\text{eff}}(f,p)$ of synaptic transmitter dynamics (equation C.2). This is the essential difference between these two synaptic parameters. Increasing the release probability increases not only the peak amplitude of the transient current but it also makes the transmitter dynamics faster (compare the width of the transient peaks on figure 5.2 top and middle panel). This has two important consequences.

First, at large release probabilities the time window for synaptic summation decreases and the neuron becomes sensitive only to *coherent* transients in the presynaptic activity. In this case, a cortical cell detects synchrony (Senn et al., 1998) among the firing of its presynaptic geniculate cells. There is good reason to expect that novel information about the outside world is contained in coherent transients in the converging inputs to a simple cell—induced by saccadic eye movements or moving objects' boundaries (c.f. Mechler et al., 1998). We conclude, that a cortical cell can extract these signal related coherent events more efficiently from the uncorrelated background noise if the transmitter release probability is high at its feed-forward synapses. This makes a neuron more sensitive to signals in a low contrast environment where the signal to noise ratio is low.

Second, the release probability determines the different Fourier components of the synaptic current in a diverse manner. The average current and higher frequency components scale with the release probability, while lower frequency components stay intact or may even decrease. In section 5.4.4 we explore this phenomenon further in the context of the experimental observation that shows differential adaptation of the average (DC) and the stimulus modulated (F1) component of the membrane potential (Carandini and Ferster, 1997). Also the fact that the release probability modifies the effective time constant $\tau_{\text{eff}}(f,p)$ of the transmitter dynamics will give a possible explanation of the phase retardation after adaptation to high contrast (Saul, 1995).

## 5.4 Simulations of contrast response and contrast adaptation—numerical results

We continue with simulations to test the main hypothesis of this paper that slow adaptation of the transmitter release probability (equation 5.10) can account for contrast adaptation. The network model is described in section 5.2 and includes dynamic processes on three different time scales. $(i)$ The membrane potential changes fast with a time constant of $\tau_{\text{m}} = 16\,\text{ms}$ (equation 5.1). $(ii)$ Fast synaptic depression due to the depletion of the synaptic transmitter is described by equation 5.3. Its effective time constant $\tau_{\text{eff}}(f,p) = \tau_{\text{rec}}/(\tau_{\text{rec}} f p + 1)$ (equation C.2) lies in the range of 15 to 200 ms depending on the presynaptic firing rate and the release probability. $(iii)$ The adaptation of the transmitter release probability (equation 5.10) is slow with a time constant of $\tau_{\text{adapt}} = 7\,\text{sec}$ (Ohzawa et al., 1985).

As a first step, we explore the effect of *fast* synaptic depression at the feed-forward and recurrent connections on the contrast response function in the model. Then, by simulating the *slow* adaptation rule for the transmitter release probability, we demonstrate that indeed our model can explain the available experimental data describing contrast adaptation. The model predicts that slow adaptation at the geniculo-cortical connections is sufficient to explain contrast adaptation, while adapting the release probability at recurrent excitatory connections changes the temporal structure of the cortical activity without significantly altering the contrast response function.

### 5.4.1 Simulation protocol, data analysis

In the simulations, the contrast response functions of the model cortical units are determined according to the experimental stimulation protocol of Carandini and Ferster (1997). The contrast response functions are calculated using an initial adaptation period and a subsequent series of interleaved test and re-adaptation periods. Each adaptation period was long enough that the release probabilities at the geniculo-cortical synapses converged to their steady state. The model network represents a small patch (an orientation-column) in the visual cortex and thus the cortical units' receptive fields are assumed to be perfectly overlapping. We simulate a drifting sinusoidal grating stimulus by modulating the firing rate of the geniculate input sinusoidally with identical phase for all cortical units. Following the original experimental paradigm, the temporal frequency $r$ of the stimulus is 2 Hz. The geniculate firing is assumed to be a Poisson process with its firing probability $p_{\mathrm{LGN}}$ as a function of time given by

$$p_{\mathrm{LGN}}(t) = \Delta t \cdot \left[ f_{\mathrm{Mod}} (\sin(2\pi r t) + m) \log_{100}(c+1) + B \right]_+ , \tag{5.13}$$

where $f_{\mathrm{Mod}} = 50\,\mathrm{Hz}$ is the maximal modulation amplitude of the firing rate; $m = 0.2$ describes the small increase of the mean firing rate with increasing contrast; $c$ is the stimulus contrast in percent; $B = 20\,\mathrm{Hz} \pm 5\,\mathrm{Hz}$ is the background activity drawn randomly for each geniculate spike train from a Gaussian distribution with a mean of 20 Hz and a standard deviation 5 Hz and kept constant during the simulation; $\Delta t = 0.2\,\mathrm{ms}$ is the simulation time step that is small enough to keep the firing probability smaller than 1. These parameters are determined such that the contrast response function of the geniculate neurons (Fig. 5.3a) resemble the experimental data (Kaplan et al., 1987). The Poisson distributed geniculate spike trains are determined for every geniculo-cortical synapse independently.

To study contrast adaptation in the model, we calculate the contrast response function for different adaptation states. The average cortical response is determined for contrast levels of 1%, 2%, 4%, 8%, 32%, 100% and the resulting data points are connected by lines. Different aspects of the cortical response as a function of contrast are calculated from the simulated cortical dynamics similar to the standard experimental methods. $(i)$ The time average of the subthreshold membrane potential and the firing rate for one stimulus contrast presentation period (DC component); $(ii)$ the stimulus frequency locked modulation of the subthreshold membrane potential and the firing rate (F1 component; the square root of the power at the temporal frequency of the stimulus) and their phase. The F1 component and its phase are calculated by performing a discrete Fourier transformation on the membrane potential and the firing rate as a function of time, where negative phase values mean delayed responses. The subthreshold membrane potential is calculated by numerically solving equation 5.1 without spikes and without a reset after spikes. The firing rate in time (defined here as the ratio of the number of firing neurons in a time step of $\Delta t$ that is divided by $\Delta t$) is calculated according to the integrate-and-fire model. All contrast response functions are averages of four independent simulations with different realizations of the same Poisson process for the geniculate firing. In figures 5.4a, 5.5a and 5.7a the time evolution is filtered by a low-pass filter with a cutoff frequency at 10 Hz.

### 5.4.2 The contrast response function

The contrast response function of cortical neurons saturates faster than their geniculate input (Albrecht and Hamilton, 1982) and the response delay decreases with stimulus contrast (e.g. Carandini, Heeger

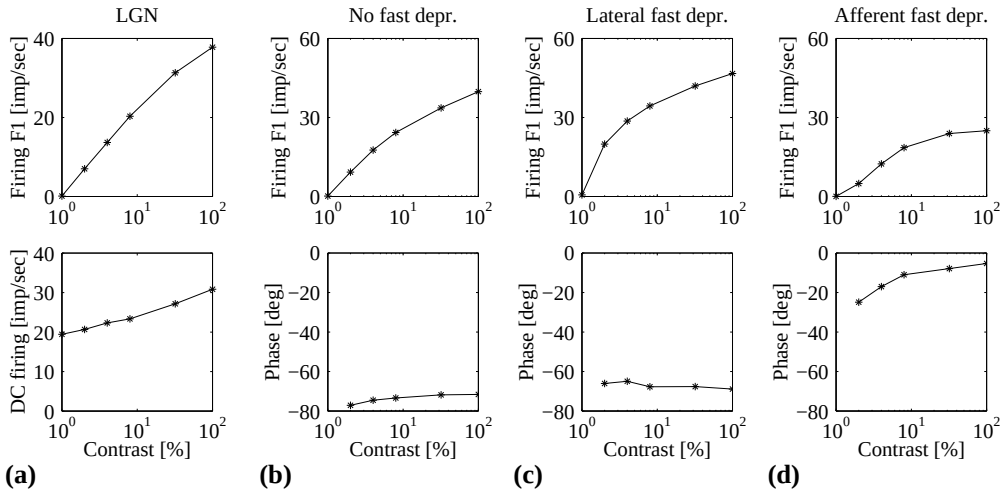


Figure 5.3: **(a)** The F1 component (top) and the DC component (bottom) of the geniculate input to the cortical layer. The F1 component of the firing rate increases linearly with log-contrast, the DC component changes only slightly. **(b,c,d)** The F1 component (top) of the cortical response and its phase (bottom) as a function of stimulus contrast. **(b)** Fast synaptic depression is present neither at the geniculo-cortical nor at the recurrent synapses. The contrast response function is almost linear with log contrast, the response phase is constant. **(c)** Fast synaptic depression is present only at the recurrent synapses. The contrast response function saturates for high contrast stimuli, but the response phase stays constant with increasing contrast. **(d)** Fast synaptic depression is present only at the geniculo-cortical synapses. The contrast response function saturates, and the response phase advances with contrast. To avoid runaway excitation and regimes where the response would saturate because of the refractoriness of neurons, synaptic weights are rescaled at synapses where fast synaptic depression is omitted. The response phase is calculated relative to the maximal phase in the presented simulations (Fig. 5.4d).

and Movshon, 1997). Here we investigate whether fast synaptic depression itself can account for these phenomena.

Figures 5.3b–d show the F1 component (top) of the cortical response in the model and its phase (bottom) as a function of stimulus contrast. If short term synaptic depression is present neither at the geniculo-cortical nor at the recurrent synapses (Fig. 5.3b), the contrast response function is almost linear as a function of log-contrast and the response phase is constant. If fast synaptic depression is present either at the recurrent excitatory (Fig. 5.3c) or the geniculo-cortical (Fig. 5.3d) synapses, the contrast response function saturates due to the saturation of the synaptic transmission (Abbott et al., 1997; Tsodyks and Markram, 1997). In accordance with the experimental data (e.g. Carandini, Heeger and Movshon, 1997) and with another modeling study (Chance et al., 1998) the model predicts a phase advance of the cortical response with contrast, but only if the geniculo-cortical synapses exhibit fast depression (Fig. 5.3d). Fast synaptic depression at the recurrent connections has no effect on the response phase (Fig. 5.3c) because the feed-forward connections trigger the cortical response and thus determine the response delay. The phase of the cortical response is contrast dependent

in the model because the effective time constant $\tau_{\text{eff}}(f,p)$ for the transmitter dynamics decreases (synaptic depression becomes faster) with increasing geniculate firing rate and thus with increasing contrast. More intuitively, at high contrast the geniculate firing rate is modulated with higher amplitude. During the dark phase of the sinusoidal grating the ON-center geniculate cells are quiet and there is time for the recovery of the transmitter. This is followed by an onset of strong geniculate activity that results in a high transient postsynaptic current and a fast depletion of the vesicles. Thus a strong rise of synaptic transmission is followed by a fast decay due to synaptic depression. The maximal synaptic transmission occurs earlier in time as stimulus contrast grows and the phase of the cortical response advances with contrast. We conclude that short term synaptic depression on the feed-forward synapses is sufficient to explain contrast saturation and a phase advance with increasing contrast, while fast plasticity of the recurrent excitatory connections alone can only account for the response saturation.

### 5.4.3 Adaptation of the geniculo-cortical synapses

In this section we present simulation results for the slow adaptation rule for the transmitter release probability at the geniculo-cortical synapses (equation 5.10). The results are compared with the experimental data that demonstrate the effects of contrast adaptation on the dynamics of the cortical response. We investigate the qualitatively different adaptation of the DC and F1 components of the membrane potential (Carandini, Barlow, O'Keefe, Poirson and Movshon, 1997) and the phase retardation by high contrast adaptation (Saul, 1995). In section 5.4.4 and 5.4.5 we focus on the role of recurrent excitation.

During the presentation of high contrast stimuli the synaptic transmitter release probability decreases two- to three-fold according to the adaptation rule. Consequently, the input/output relation of the model neurons also changes. Figure 5.4 shows the response of the cortical units in the model after adaptation to 1% (solid lines) and to 50% (dashed line) contrast. The simulations involved strong recurrent excitatory amplification and slow adaptation of the release probability of the geniculo-cortical synapses. The adaptation of the F1 component of the subthreshold membrane potential is different from the adaptation of the average (DC) subthreshold membrane potential. This effect will be further studied in the next section. The model predicts a 2-5 mV change in the DC subthreshold membrane potential. This shift is smaller than was observed experimentally (Carandini, Barlow, O'Keefe, Poirson and Movshon, 1997). Note that in the original experiment (Carandini, Barlow, O'Keefe, Poirson and Movshon, 1997) the subthreshold synaptic modulation of the membrane potential could not be measured directly and was instead determined indirectly by removing the spikes. The ambiguity in this process could explain the observed discrepancies in the experimental data and the simulation results: the average subthreshold membrane potential could be overestimated proportionally to the firing rate if segments of the supra-threshold parts of the spikes are not removed. The F1 component of the firing (Fig. 5.4c) adapts strongly because the average membrane potential decreases and a smaller part of the membrane potential remains over the firing threshold.

In accordance with the experimental data obtained for simple cells in cat V1 (Saul, 1995), the delay (defined as the phase of the F1 component) of the cortical response increases in the model after adaptation to high contrast. After adaptation to high contrast, the transmitter release probability $p$ at the geniculo-cortical connections decreases and the synaptic depression becomes slower, because its effective time constant $\tau_{\text{eff}}(f,p)$ scales with $1/p$. In contrast, after adaptation to low contrast, when the transmitter release probability is high, the synaptic depression becomes faster: a sudden rising phase of the synaptic current at the geniculo-cortical synapses is rapidly followed by fast depression.

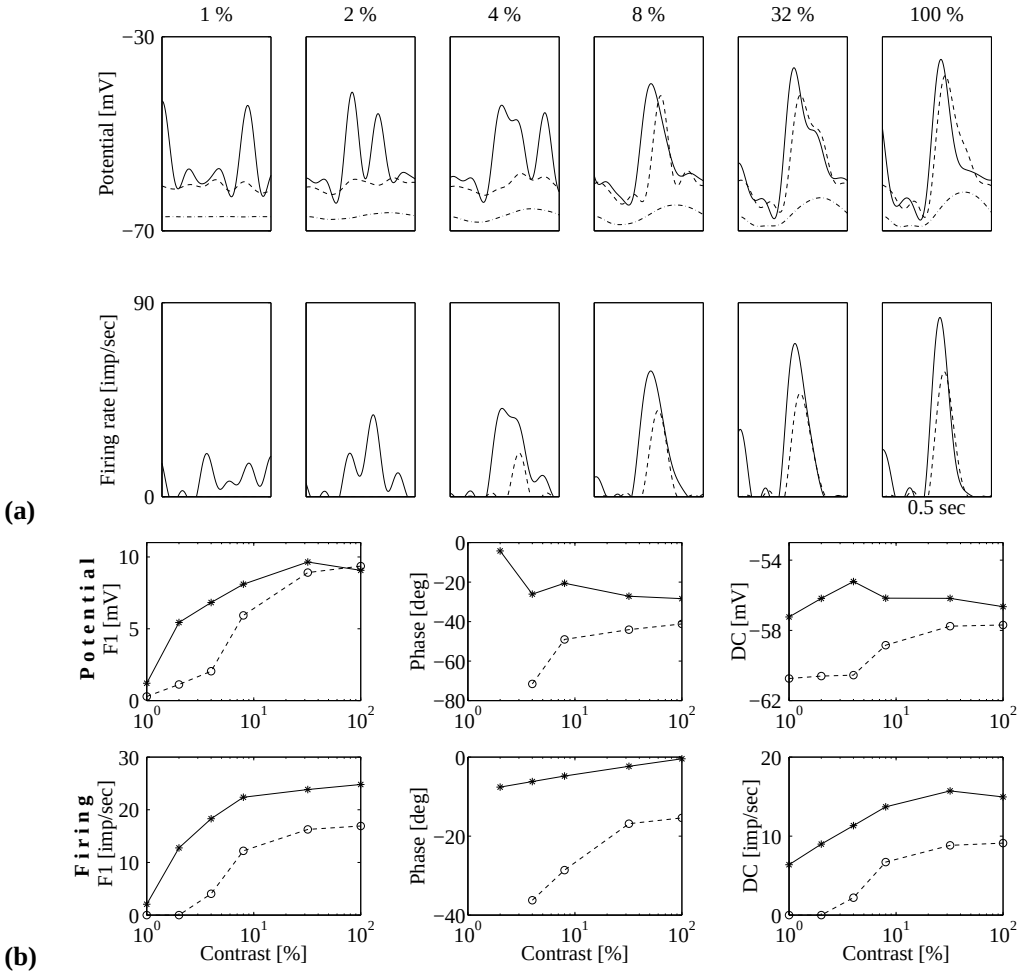


Figure 5.4: Simulation results *with* strong recurrent excitatory connections and with slow adaptation of the release probability of the geniculo-cortical synapses; the release probability of the recurrent synapses is fixed to $p = 0.4$. (**a**) Cortical response in the model to a sinusoidally modulated input (dashed-dotted line, the curve is scaled to fit on the plot and shows the geniculate firing rate). (**b**) The F1 component, the phase of the F1 component, and the average of the subthreshold membrane potential (top) and the firing (bottom) are plotted as a function of stimulus contrast after adaptation to 1% (solid lines) and to 50% (dashed lines) contrast stimuli.

The maximal cortical firing occurs earlier in time, the phase of the F1 component of the cortical firing increases after low contrast adaptation. This observation indicates that contrast adaptation cannot be the sole consequence of a decrease in synaptic weights or an increase of the firing threshold—e.g. induced by a tonic hyperpolarization. All of these effects would not alter the phase of the cortical response.

The experimental data suggest that the response amplitude and phase are not coupled to each

other (Saul, 1995). The phases of cortical responses with the same F1 component can differ after adaptation to different contrast levels. Our model accounts for this observation (Fig. 5.4b). As highlighted above, different factors make the response phase dependent on the stimulus contrast and on the contrast adaptation. In the first case, the response phase changes because the effective time constant of the transmitter dynamics $\tau_{\text{eff}}(f,p)$ is affected by the contrast dependent presynaptic firing rate; in the second case, $\tau_{\text{eff}}(f,p)$ is modulated by the adaptation of the release probability $p$—the two factors determine the neural response and its phase in a different manner.

### 5.4.4 Recurrent excitation and contrast adaptation

Figure 5.5 shows simulation results for contrast adaptation with the weights of the recurrent excitatory connections set to zero. Without recurrent excitation, the F1 component of the membrane potential is independent of contrast adaptation similarly to the experimental data (Carandini and Ferster, 1997). Modulating the transmitter release probability does not have a simple multiplicative effect on the postsynaptic response, as might be expected. Both the peak amplitude of the excitatory postsynaptic current and the rate of the transmitter dynamics scales with the release probability ($\tau_{\text{eff}}(f,p)$ decreases with increasing the release probability, see sections 5.3.2 and C.1). After adaptation to low contrast the release probability is large, and after a fast rise the synaptic current decreases rapidly because the synaptic depression is faster. After adaptation to high contrast the release probability is small, and the peak current is lower but the synaptic depression is slower too. Changing the amplitude of the current is compensated by increasing its width in time. It follows that the low frequency components of the membrane potential do not change with the release probability, while the high frequency components scale with the release probability (compare the solid and dashed lines in figure 5.6). It depends on the temporal frequency of the stimulus whether the F1 component of the membrane potential changes strongly or not. The model predicts stronger adaptation of the F1 component of the membrane potential if the temporal frequency of the stimulus is increased. Adaptation of the release probability of the feed-forward synapses affects the DC component of the membrane potential in the absence of recurrent excitation. With modulating the DC membrane potential, smaller or larger part of the membrane potential remains over the firing threshold and the F1 component of the firing rate adapts even if there is no adaptation in the F1 component of the membrane potential. If strong recurrent excitation is present (as in section 5.4.3) the adaptation in the F1 component of the cortical firing is projected back to the cortical neurons modulating the F1 component of the membrane potential. Based on these predictions we can conclude that for a stimulus with low temporal frequency, the contrast adaptation of the F1 component of the membrane potential scales with the strength of the recurrent excitation. The adaptation of the F1 component of the firing rate could therefore be used to measure the effective strength of the recurrent excitatory input to a simple cell in V1.

### 5.4.5 Modifying the release probability of the recurrent excitatory synapses

In a next step, we study whether contrast adaptation can be evoked by changing the transmitter release probability of the recurrent excitatory synapses alone (figure 5.7). The model predicts that the contrast response function is largely independent of the release probability at the recurrent excitatory connections. Interestingly, we find that increasing the release probability (from 0.2 to 0.7) of the recurrent excitatory synapses—in contrast to the geniculo-cortical synapses—does not alter or may

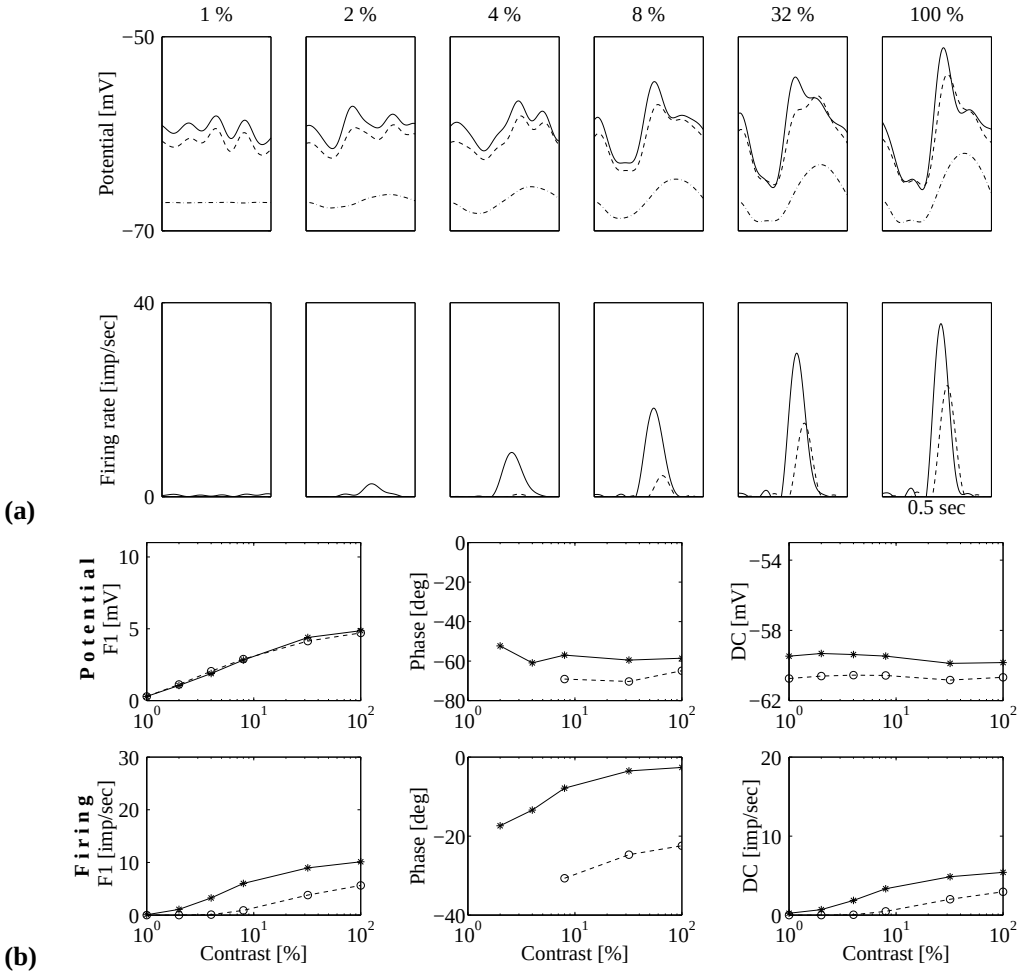


Figure 5.5: Simulation results *without* recurrent excitation and with slow adaptation of the release probability of the geniculo-cortical synapses. (**a**) Cortical response in the model to a sinusoidally modulated input (dashed-dotted line, the curve is scaled to fit on the plot and shows the geniculate firing rate). (**b**) The F1 component, the phase of the F1 component, and the average of the subthreshold membrane potential (top) and the firing (bottom) are plotted as a function of stimulus contrast after adaptation to 1% (solid lines) and to 50% (dashed lines) contrast stimuli.

slightly *decrease* the F1 component of the firing rate of the cortical units. In that case, the model predicts a change in the time structure of the cortical activity, as we will show.

In Fig. 5.8a the geniculate (dashed-dotted line) and cortical (solid and dashed lines) firing probability for 30 model cortical units at high (solid line) and at low (dashed line) transmitter release probability are plotted as a function of time for 100% stimulus contrast. The autocorrelation functions of the cortical activity are shown in figure 5.8b. If the release probability is high at the recurrent excitatory connections, synchronized oscillations emerge at approximately 50 Hz. Note

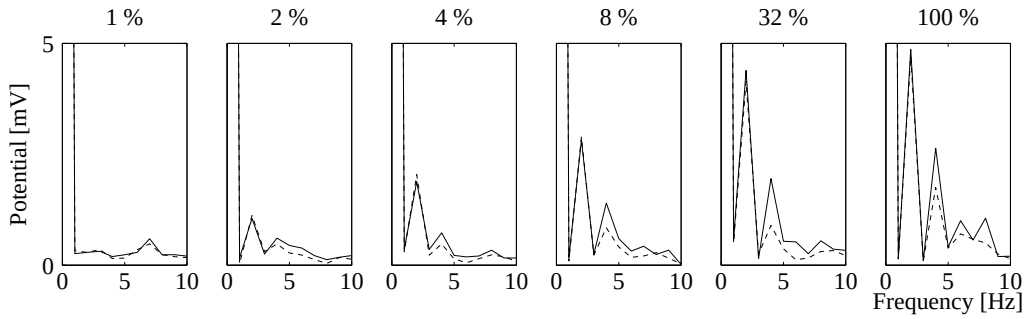


Figure 5.6: The Fourier-spectra of the membrane potential for the model configuration *without* recurrent excitation (the same simulation as in figure 5.5) after adaptation to 1% (solid lines) and to 50% (dashed lines) contrast stimuli.

that the oscillation remains stable even in the presence of strong, Poisson distributed irregular geniculate input. Oscillatory activity emerges because of the following reasons. Increasing the release probability of the recurrent connections leads to a more reliable synaptic transmission and to larger EPSP amplitudes. Active cells initiate a burst of highly synchronous firing in the cortical layer via the strong recurrent connections. As a trade off, at high release probability, most of the available synaptic transmitter is released after a presynaptic firing event such that efficient recurrent excitation can occur only after some recovery time for the transmitter. Thus after a period of high cortical activity, the recurrent reamplification fades away because only a small amount of transmitter remains available. This leads to oscillatory behavior. The frequency of oscillation depends on the effective time constant $\tau_{\text{eff}}(f,p)$ of the transmitter dynamics. The phenomenon of emergent oscillations due to depressing synapses in a purely excitatory network has been explored by Senn et al. (1996) and it is similar in nature to emergent oscillations due to slow adapting currents (Crook, Ermentrout and Bower, 1998) or delayed inhibition (e.g. König and Schillen, 1991). For low synaptic transmitter release probability, the autocorrelogram is flatter and oscillatory firing is not as prominent any more. The frequency of the oscillation decreases because the effective time constant $\tau_{\text{eff}}(f,p)$ for the synaptic transmission grows with decreasing the release probability. In parallel, the F1 component of the cortical firing becomes slightly stronger because the synaptic transmitter is utilized in a more "economic" way.

We conclude that adaptation of the transmitter release probability of the *geniculo-cortical* synapses is sufficient for contrast adaptation. In contrast, plasticity of the release probability at the recurrent excitatory connections cannot induce adaptation of the contrast response function. This model prediction differs from predictions of a previous modeling study (Todorov, Siapas, Somers and Nelson, 1997) that assumed plasticity of the *strength* of both the geniculo-cortical and the recurrent excitatory synapses and also recurrent inhibition to achieve contrast adaptation. Our simulations reveal that modifying the transmitter release probability of the recurrent excitatory connections changes the time structure of the cortical response. Increasing the release probability of the recurrent excitatory synapses leads to oscillatory activity. These results suggest an efficient functional segregation of synaptic plasticity at the geniculo-cortical and the recurrent excitatory connections. Adapting the release probability at the geniculo-cortical synapses may play a key role in contrast adaptation, while—without affecting the contrast response function—increasing the release proba-

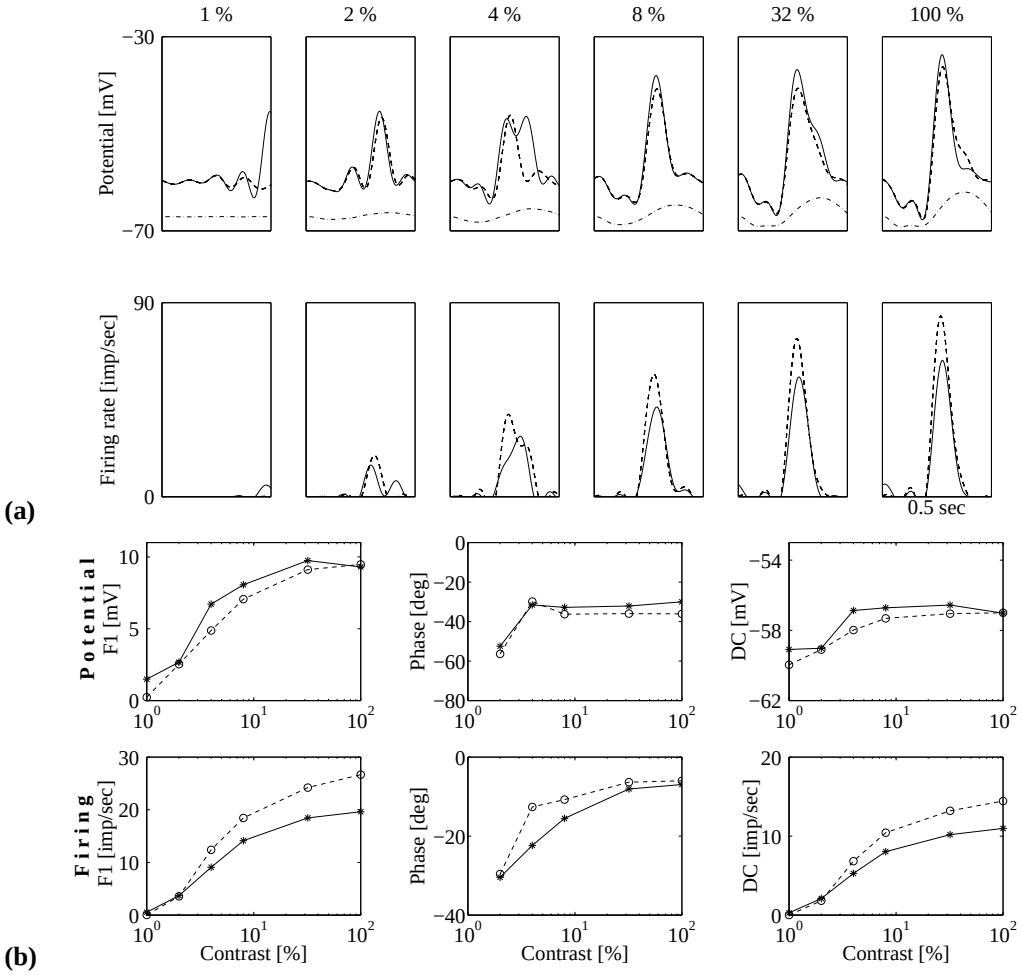


Figure 5.7: Simulation results *with* strong recurrent excitation. The release probability of the recurrent excitatory connections is set to 0.7 (solid lines) and 0.2 (dashed lines). Slow adaptation of the release probability of the geniculo-cortical synapses is *not* present, the release probability is fixed to 0.5. (**a**) Cortical response in the model to a sinusoidally modulated input (dashed-dotted line, the curve is scaled to fit on the plot and shows the geniculate firing rate). (**b**) The F1 component, the phase of the F1 component, and the average of the subthreshold membrane potential (top) and the firing (bottom) are plotted as a function of stimulus contrast.

bility can shift the cortical network to a regime of prominent oscillatory activity. Modulation of the release probability of the recurrent excitatory synapses could thus play a key role in dynamic feature binding and segregation in the visual cortex (e.g. von der Malsburg and Buhmann, 1992; Engel et al., 1997).

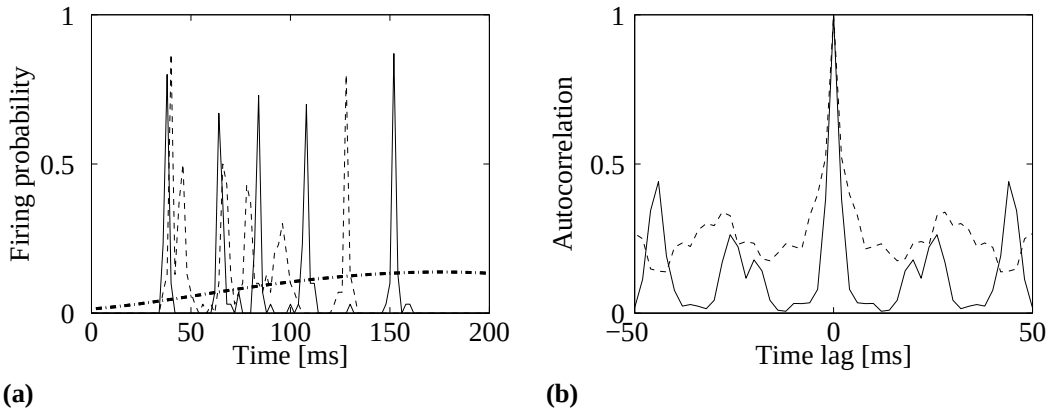


Figure 5.8: **(a)** The cortical response of the model (not low-pass filtered in this case) if the transmitter release probability is changed only at the recurrent excitatory connections (the same simulation as in figure 5.7). The average firing probability of 30 model cortical units is calculated for 2 msec time bins at high (solid line) and at low (dashed line) release probabilities. The geniculate activity (modulated sinusoidally with 2 Hz, dashed-dotted line) is in its rising phase and corresponds to 100% contrast stimulus. **(b)** The autocorrelation function of the firing probability at high (solid line) and at low (dashed line) release probability. The autocorrelation is normalized to 1.

## 5.5 Possible physiological indications of the transmitter release probability adaptation

In this section we endetail the model predictions to support experimental tests of our hypothesis. Specifically we focus on possible physiological effects of changing the transmitter release probability. We simulate membrane potential dynamics and synaptic behavior based on the model described in section 5.2.2. This computational model is really simplified compared to the complexity of a single synapse or a dendrite. It follows, that these simulations describe general tendencies, and they cannot be interpreted literally. Probably the most important factors, that the model lacks are the mechanisms that underly fast synaptic facilitation, and active dendritic summation. These mechanisms are not included in the model, because the currently available physiological data are insufficient.

The model predicts, that the geniculo-cortical synapses are involved in contrast adaptation (see sections 5.4.3, 5.4.5). The experimental setup should involve full cell recording in V1 simple cells and electrical stimulation of the geniculate axons that innervate the recorded cell. This setup could make it possible to explore geniculo-cortical synapses *in vivo*. The EPSP trains evoked by certain geniculate stimulation patterns should be recorded after adaptation to different contrast levels. The experiment, therefore, involves a long adaptation period with visual stimulus followed by a test period to estimate the transmitter release probability on the geniculo-cortical synapses that connect the stimulated axon with the recorded cell. Naturally, one has to be careful to minimize adaptation by the testing presynaptic electrical pulses.

The number of free parameters should be limited if it is possible. In our description there are three important parameters of the synapses: the recovery time constant $\tau_{rec}$, the transmitter release

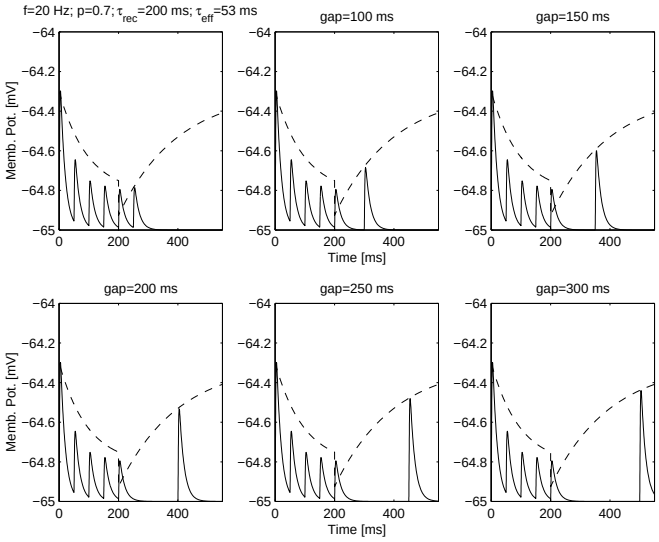


Figure 5.9: Determining the recovery time constant $\tau_{rec}$ of the transmitter. The simulations illustrate the experimental protocol and the EPSP trains predicted by the model (solid line). The dynamics of the mean EPSP averaged over different realizations of a Poisson distributed presynaptic spike train with a frequency of 20 Hz is represented by the dashed line.

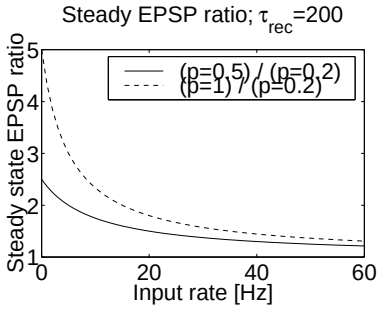


Figure 5.10: The relative increase of the steady state EPSP amplitude if the release probability is increased from p=0.2 to p=0.5 (solid line) and to p=1.0 (dashed line). Note that the difference is more pronounced at *low* presynaptic rates.

probability $p$, and maximal synaptic conductance $g_{max}$. This latter parameter scales the strength of the synaptic transmission without modifying its dynamic behavior. In contrast, the recovery time constant and the release probability both modifies the dynamics of the the fast synaptic depression. These two parameters should determined independently.

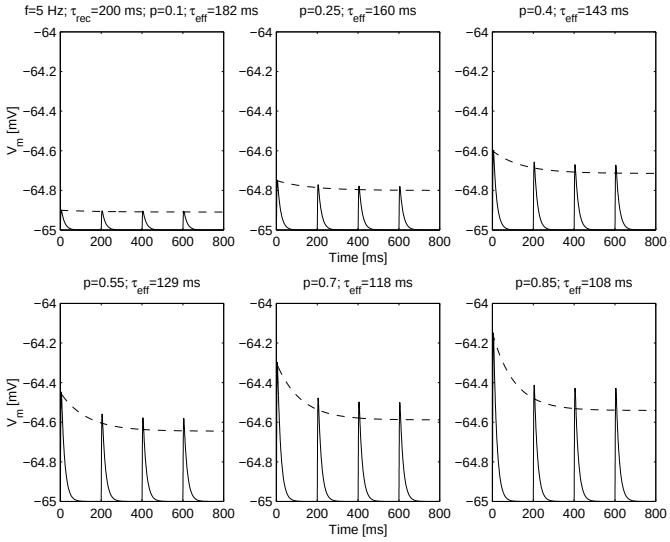


Figure 5.11: The membrane potential dynamics (solid line) for different release probabilities $p$ at a presynaptic stimulation of 5 Hz. The expected membrane potential for Poisson distributed presynaptic spike trains with the same mean is depicted by the dashed line.

### 5.5.1 Determining the recovery time constant $\tau_{\text{rec}}$

The time course of the transmitter recovery could be explored by the following stimulation protocol:

1. Apply a high frequency (10-20 Hz) presynaptic stimulation for a long enough period ($\Delta T$) that the available amount of transmitter is depleted to its steady state level ($\Delta T \approx 3\tau_{\text{eff}}(f,p)$).

2. After this period of presynaptic stimulation leave a gap of silence. For different trials the duration of this gap should be increased step by step.

3. After the silence period apply a short electrical pulse that evokes an EPSP. In the case of a purely fast depressing synapse the EPSP amplitude should recover approximately exponentially with a time constant of $\tau_{\text{rec}}$. If there is facilitation, then at a certain (long enough) period of silence the amplitude of the evoked EPSP could be bigger than the very first EPSP in the train.

### 5.5.2 Determining the transmitter release probability $p$

Our model predicts that the after adaptation to high contrast, the release probability decreases. The release probability could be estimated experimentally by applying a train of three-four presynaptic pulses and investigate the rate of decrease in the EPSP amplitudes. The ratio of the steady state EPSP amplitudes in the model is shown in Fig. 5.10 if the release probability is increased from p=0.2 to p=0.5 and to p=1.0. The effect is largest at low presynaptic firing rates. However, note that at low presynaptic firing rates the EPSP amplitude converges to its steady slower. Therefore, there is trade

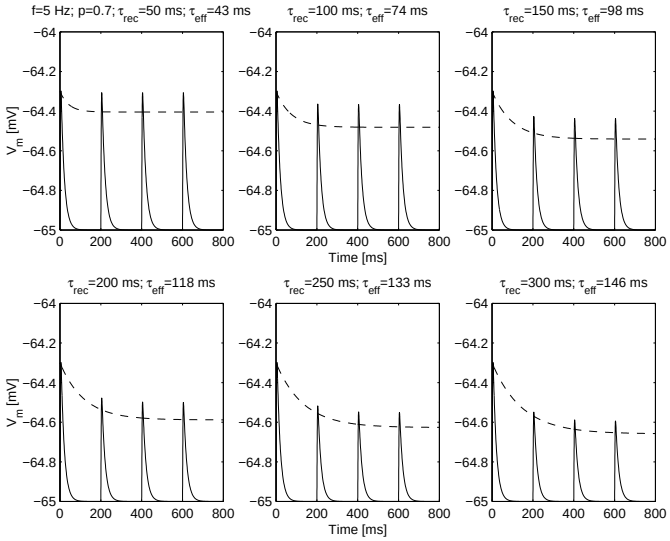


Figure 5.12: The membrane potential dynamics (solid line) for different time at a presynaptic stimulation of 5 Hz. The expected membrane potential for Poisson distributed presynaptic spike trains with the same mean is depicted by the dashed line.

off between the magnitude of the effect and the speed of convergence. Simulation results for the EPSP trains evoked by a presynaptic spike train with a frequency of 5 Hz are presented in Fig. 5.11.

Note that the rate of depression increases with both increasing the release probability and the increasing the recovery time constant of the synaptic transmitter. However there is an important difference. The amplitude of the first EPSP decreases with the release probability but it is invariant of the recovery time constant (compare Figs. 5.11 and 5.12).

## 5.6 Discussion

We have shown that fast synaptic depression of the geniculo-cortical synapses can explain the saturation of the contrast response function and the phase advance of the visual cortical response with increasing contrast. We have tested our hypothesis that a slow adaptation of the transmitter release probability at the geniculo-cortical synapses can account for contrast adaptation. Assuming a slowly changing input from the LGN, an adaptation rule was derived for the release probability to maximize mutual information between a visual cortical neuron's input and output firing rates. The synaptic dynamics for abruptly changing presynaptic firing rates was investigated separately in section 5.3.2. Results show that the peak amplitude of the transient synaptic current scales with the release probability. Increasing the release probability also increases the rate of synaptic depression. As a consequence, the temporal window for synaptic input summation decreases and a neuron becomes sensitive only to a synchronous increase in its afferent activity. This synchronous firing is likely to be signal related and therefore a cortical cell could extract the signal related component of the geniculate input more efficiently if the release probability at the geniculo-cortical connections

is high. In the subsequent numerical simulations the dynamics of the release probability obey an adaptation rule which maximizes the mutual information between the neurons' input and output in the case of slowly changing input. The simulation results indicate that the slow adaptation of the transmitter release probability of the geniculo-cortical synapses can be the underlying neural mechanism for contrast adaptation. The most important point of our investigation is that in contrast to the synaptic weight, the release probability modifies the effective time constant of the transmitter dynamics too. Based on this property of the release probability, we can explain all experimental observations indicating that contrast adaptation alters the dynamic properties of the cortical response. Changing the release probability of the recurrent excitatory synapses, however, has only a weak effect on the contrast response function. Within the model framework we found that after increasing the transmitter release probability of the recurrent excitatory connections, synchronous oscillatory activity emerges. We therefore propose that slow adaptation of the transmitter release probability of the geniculo-cortical and the recurrent synapses have independent functional roles in the primary visual cortex.

### 5.6.1 Model predictions

The good agreement between our simulation results and the currently available data gives an indirect verification for our hypothesis that changing the transmitter release probability or, in other words, redistribution of the available amount of synaptic transmitter accounts for contrast adaptation. The modulation of the transmitter release probability changes the frequency dependent fast synaptic depression. Our hypothesis, that adaptation of the transmitter release probability accounts for contrast adaptation, could be directly justified experimentally by intracellular measurements of EPSP series evoked by stimulating the geniculo-cortical axons. Possible experimental steps illustrated with the model predictions are described in detail in section 5.5. An alternative and simpler physiological indication of the modulation of the transmitter release probability could be the paired pulse facilitation ratio (PPF). The model predicts that after adaptation to low contrast, the release probability increases, consequently the PPF ratio decreases and vice versa.

Simulations revealed that the contrast adaptation of the F1 component of the subthreshold membrane potential depends on two factors, *(i)* the strength of the recurrent excitatory connections and *(ii)* the temporal frequency of the grating stimulus. *(i)* The adaptation of the cortical firing is fed back by the recurrent excitatory connections. Therefore the adaptation of the F1 component of the subthreshold membrane potential increases with the effective strength of the recurrent excitatory feedback. This observation could provide a new experimental technique to estimate the strength of the recurrent excitatory coupling. As several computational models predict that strong recurrent excitation is required for establishing sharp orientation tuning, it would be very interesting to explore how adaptation of the F1 component of the membrane potential and the sharpness of orientation tuning are correlated to each other. *(ii)* Adaptation of the transmitter release probability effects mainly the average and the higher frequency components of the synaptic transmission. As a natural consequence, the model predicts that the F1 component of the membrane potential adapts stronger for high temporal frequency grating due to the synaptic adaptation at the geniculo-cortical synapses.

The model predicts that the slow adaptation of synaptic transmission is local to the presynaptic side of the geniculo-cortical synapses. Contrast adaptation thus the adaptation depends on the firing rate of the presynaptic geniculate cells targeting a simple cell in V1. Induced by adapting stimuli with different orientations, the reduction of the short latency EPSPs—presumably of geniculate origin—could therefore constitute the geniculate origin orientation bias to a simple cell. However, a model

of a hypercolumn with full recurrent circuitry should give more exact predictions in this respect.

### 5.6.2 Model assumptions

In order to investigate the hypothesis that synaptic plasticity of the excitatory connections is sufficient to explain contrast adaptation we studied a model without recurrent inhibition. Even though the simulation results account well for the phenomenology of contrast adaptation—given the controversial data regarding the quality and role of recurrent inhibition—we cannot exclude the possibility that adaptation in V1 is a cooperative phenomenon which involves modulation by recurrent inhibition of network origin as speculated by (Vidyasagar, 1990; Ahmed et al., 1997). But even under the assumption that recurrent inhibition plays a role in contrast adaptation or lateral gain control, the neural mechanism that modulates inhibition remains unresolved. Based on the available experimental data and our computational study we suggest that plasticity of the geniculo-cortical synapses could be a good candidate.

For the derivation of the adaptation rule (Eq. 5.10) we used the steady state approximation for the synaptic transmitter (Eq. C.3). Our investigation on the dynamics of depressing synapses in the presence of abruptly changing input (section 5.3.2) showed that at faster modulation of the presynaptic firing rate the transient postsynaptic current is significantly different from the steady state current. Thus for studying the cortical response to more realistic stimuli—in which transient changes carry most of the novel information—it is essential to incorporate the dynamic behavior of the synapses into the learning rule. An extended learning rule would account for the transients in the EPSC induced by the sharp moving edges in the visual environment and by the saccadic eye movements.

Exploring contrast adaptation in the visual cortex we focused on synaptic mechanisms and derived an adaptation rule for a presynaptic parameter with the objective to maximize the mutual information between a neuron's input and output. In agreement with experimental studies (DeBruyn and Bonds, 1986; Vidyasagar, 1990) we assumed that adaptation of active ion channels is not involved in contrast adaptation and we kept the single cell model as simple as possible using the integrate-and-fire model framework. However, on a slower time scale and in different regions of the nervous system, adaptation of active channel properties could also be a powerful mechanism for maximizing the information content of a neuron's output about its input (Stemmler and Koch, 1999). It may be possible that adaptation of synaptic transmission and active channel properties act in a cooperative manner on different time scales.

### 5.6.3 Contrast adaptation and the receptive field profile

In this paper we proposed that contrast adaptation serves for maximizing the mutual information between a single V1 neuron's input and output. The global image contrast in the visual environment changes on the scale of several seconds and therefore contrast adaptation should follow this time scale too. We argued that an adaptation mechanism with a time constant of 7 sec would make it possible to estimate the global image contrast, that is mainly reflected in the variance of the firing rates of the geniculate cells calculated for several subsequent fixation periods. We also proposed that plasticity of the synaptic transmitter release probability accounts for contrast adaptation. It is widely accepted that the development of receptive field structures of neurons is also related to synaptic plasticity, but on a much longer time scale. Receptive field profiles are thought to be determined by the higher order moments (than the variance) of the visual world statistics that express patterns

or structure. This higher order structure is rather static and its estimation also requires longer time, therefore receptive field profiles should be determined and constrained by another less plastic synaptic parameter than the release probability. One possible candidate for this would be the postsynaptic maximal conductance $g_{\max}$.

# Appendix A

# Parameters for the feed-forward model in chapter 2

Some of the measured properties of the geniculate cells cannot be applied directly in the model. Here we calculate some parameters based on the experimentally measured data.

## A.1 Receptive-field parameters of the LGN M cells in the model

Here we calculate the parameters for the receptive fields used in the rate model for the intracortical generation of orientation selectivity (chapter 2). Data from (Spear et al., 1994) and (Croner and Kaplan, 1995) showed that $r_c$ and $r_s$ depend roughly linearly on eccentricity, hence the value $r$ at 5° eccentricity is approximately equal to the average $\bar{r}$ over the interval $0° - 10°$ provided by (Spear et al., 1994).

(Croner and Kaplan, 1995) showed for retinal ganglion cells that the integrated sensitivity of the ON and OFF subfields is constant regardless of the eccentricity, i.e. the peak sensitivity $k$ is approximately proportional to $1/r^2$. Hence peak sensitivities $k_{c,s}$ can be obtained via

$$k = a/r^2. \tag{A.1}$$

The constant $a$ can be calculated from the average value $\bar{k}$

$$\log(\bar{k}) = \frac{1}{r(0°) - r(10°)} \int_{r(0°)}^{r(10°)} dr \log(\frac{a}{r^2}) , \tag{A.2}$$

which is given in (Spear et al., 1994).

## A.2 Parameters of the Geniculate Transfer Function

The parameters of the transfer function $g_{\mathrm{M}}(R)$, eq. 2.3, were determined as follows. The threshold $T_M$ was set equal to the afferent activity which a geniculate M-cell receives for a grating stimulus of

optimal spatial frequency at the contrast threshold $c_{\mathrm{M}}^{\min}$:

$$T_{\mathrm{M}} = k_{\mathrm{c}}\pi r_c^2 - k_{\mathrm{s}}\pi r_s^2 + c_{\mathrm{M}}^{\min} F(f_{\mathrm{opt}})\,, \tag{A.3}$$

where

$$F(f_{\mathrm{opt}}) = k_{\mathrm{c}}\pi r_c^2 \exp(-(\pi f_{\mathrm{opt}} r_{\mathrm{c}})^2) - k_{\mathrm{s}}\pi r_s^2 \exp(-(\pi f_{\mathrm{opt}} r_{\mathrm{s}})^2)\,. \tag{A.4}$$

In order to determine $b$ we assume that a grating stimulus is presented whose positive half is centered on the cell's receptive-field. The transfer function, eq. (2.3), can then be rewritten as a function of the contrast $c$ of the grating,

$$g_{\mathrm{M}}(c) = \begin{cases} 0 & \text{if } c \le c_{\mathrm{M}}^{\min} \\ \frac{c - c_{\mathrm{M}}^{\min}}{c - c_{\mathrm{M}}^{\min} + \frac{b}{F(f_{\mathrm{opt}})}} & \text{otherwise}\,, \end{cases} \tag{A.5}$$

from which we can obtain $b$ via the averaged contrast gain normalized to the maximal geniculate activity $M_{\max}$

$$\frac{M_{\max}}{G} = \frac{b}{F(f_{\mathrm{opt}})}\,. \tag{A.6}$$

# Appendix B

# Empirical entropy manipulation—derivations for chapter 4

Here we describe the numerical procedure used in chapter 4 to estimate and maximize mutual information w. r. t. input-output mapping parameters. We consider two models for the output noise: additive Gaussian and multiplicative Poisson. For the additive output noise case, following (Viola et al., 1996), we obtain a Parzen-estimation for the empirical entropy and we maximize the mutual information based on this estimate. For the Poisson-spiking model, the output entropy is estimated directly from the spike count histograms.

## B.1 Estimating the mutual information

In this section some basic properties of the mutual information is recalled. We consider the following stochastic input-output relation

$$\vec{y} = g(\vec{x}) + n = \vec{u} + n(\vec{u}) \,, \tag{B.1}$$

where the mapping $g : X \to Y$ is assumed to be deterministic (not necessarily invertible) and $n(\vec{x})$ is an additive/multiplicative noise process. The input variable $\vec{x}$ can be interpreted as the feed-forward input to cortical neurons, the mapping $g(\vec{x})$ refers to the transformation of the input by the recurrent network and $n(\vec{x})$ is the inherent noise on the cortical neurons' activity. The task is to estimate the mutual information between input vector $X$ and the output vector $Y$ of the network. For this, first, we need to consider an important theorem about the mutual information.

It has been shown (Nadal et al., 1998) that if the mapping is deterministic, then the mutual information between the signal $X$ and the neural output $Y$ is the same as the mutual information between the deterministic output $U$ and the stochastic output $Y$:

$$I[X,Y] = I[U,Y] \,. \tag{B.2}$$

The mutual information between the continuous variables $X$ and $Y$ is

$$\begin{aligned} I(X,Y) &= \int d\vec{x}\, p(\vec{x}) \int d\vec{y}\, p(\vec{y}|\vec{x}) \ln \frac{p(\vec{y}|\vec{x})}{p(\vec{y})} \\ &= \int d\vec{x}\, p(\vec{x}) \underbrace{\int d\vec{u}\, \delta(\vec{u} - f(\vec{x}))}_{1} \int d\vec{y}\, p(\vec{y}|\vec{x}) \ln \frac{p(\vec{y}|\vec{x})}{p(\vec{y})} \,. \end{aligned} \tag{B.3}$$

Changing the integration order and noting that $p(\vec{u}) = \int d\vec{x}\, \delta(\vec{u} - f(\vec{x}))$ and $p(\vec{y}|\vec{x}) = p(\vec{y}|\vec{u})$ we get finally

$$I[X,Y] = \int d\vec{u}\, p(\vec{u}) \int d\vec{y}\, p(\vec{y}|\vec{u}) \ln \frac{p(\vec{y}|\vec{u})}{p(\vec{y})} = I(U,Y) \tag{B.4}$$

If the noise is additive, then

$$H[Y|U] = H[N] \,. \tag{B.5}$$

Because $I[X,Y] = I[U,Y] = H[Y] - H[Y|U]$ as proven above, the estimation of the mutual information $I[X,Y]$ simplifies to the estimation of the output entropy only. If the noise is not additive, then one must estimate both the entropy of the output $H[Y]$ and the conditional entropy $H[Y|U]$. This latter is clearly more complex computationally because a higher dimensional (input + output dimensions) density function has to be estimated.

## B.2 Empirical entropy manipulation—Additive noise

Here we assume that the neural code is the firing rate of the neurons, that is corrupted by an additive Gaussian noise. In this framework the output variable is continuous. Optimizing the input-output mapping $g : X \to Y$ involves to steps. First, the empirical entropy is estimated. The output probability distribution of the network is estimated by a maximum-likelihood approach by stochastic gradient ascend. The input-output mapping function $g : X \to Y$ is modified based on this estimate such that the mapping maximizes the output entropy and therefore the mutual information. Modifying the mapping function changes the entropy of the output. It follows that after a learning step for $g$, the entropy has to be re-estimated. The algorithm for the empirical entropy manipulation is then

```
loop while I*[X,Y] < max
      loop while p*(y) is not converged
            learning step to obtain p*(y) ≈ p(y)
            (gradient ascend on the model parameter likelihood)
      end
      learning step for g(x) to maximize I*[X,Y]
      (gradient ascend on the estimated entropy w.r.t. g(x))
end
```

where $p^*(y)$ is the estimate for the probability density $p(y)$ of the output and $I^*[X,Y]$ is the estimate for the mutual information based on $p^*(y)$. In the following we describe the two learning steps involved in this iterative algorithm.

### B.2.1 Parzen estimate for the empirical entropy

Here we outline a procedure for estimating the entropy of the continuous random variable $Y$. Estimating the entropy requires the estimation of the whole probability distribution function of the random variable. A solution for this generally really complex problem has been suggested by Viola et al. (1996), who introduced a Parzen estimator for the density function. Based on a sample $A \subset Y$, the density at $y \in Y$ is estimated as a sum of Gaussian basis functions $G_\Sigma$ with zero mean and covariance matrix $\Sigma$

$$p^*(y|A,\Sigma) = \frac{1}{N} \sum_{y_A \in B} G_\Sigma(y - y_A) \,. \tag{B.6}$$

The model parameters, that is the covariance matrix $\Sigma$ is obtained by maximizing the model's likelihood w. r. t. to the parameters. This is equivalent to minimizing the empirical entropy $h^*_\Sigma(A)$ given sample $A$. Using the "*one leave out cross-validation*" one can avoid the trivial solution of Gaussian kernels centered at the data points with zero variances. The empirical entropy is based on the Parzen estimation for the distribution

$$\begin{aligned} h^*_\Sigma(A) &= -\frac{1}{N_A} \sum_{y_A \in A} \log\left(p^*(y_A|A/\{y_A\},\Sigma)\right) \\ &= -\frac{1}{N_A} \sum_{y_A \in A} \log\left(\frac{1}{N_A - 1} \sum_{y'_A \in A/\{y_A\}} G_\Sigma(y_A - y'_A)\right) \\ &= -E_A\left[\log E_{A/\{y_A\}}\left[G_\Sigma(y_A - y'_A)\right]\right] \\ &= -l(A;\, A/\{y_A\},\Sigma) \,. \end{aligned} \tag{B.7}$$

where, $E_A$ is the expectation over the sample set $A$. Each Gaussian basis function has the same covariance matrix $\Sigma$ that is assumed to be diagonal with the elements of $\Sigma_{ii} = \sigma_i$. The variances of the Gaussian basis functions that account best for the real data are obtained by gradient descend on the empirical entropy (Eq. B.7). Then the learning rule for the model parameters is

$$\Delta\sigma_i \propto -\frac{\partial}{\partial \sigma_i} h^*_\Sigma[A] = E_{y_A \in A}\left[\frac{\frac{1}{N_B}\sum_{y_B \in B} G_\Sigma(y_B - y_A)\frac{1}{\sigma_i}\left(\frac{(y_B - y_A)_i^2}{\sigma_i^2} - 1\right)}{\frac{1}{N_B}\sum_{y_B \in B} G_\Sigma(y_B - y_A)}\right] \tag{B.8}$$

$$= E_{y_A \in A}\left[\sum_{y_B \in B} \frac{G_\Sigma(y_B - y_A)}{\sum_{y'_B \in B} G_\Sigma(y'_B - y_A)} \frac{1}{\sigma_i}\left(\frac{(y_B - y_A)_i^2}{\sigma_i^2} - 1\right)\right] \tag{B.9}$$

$$= E_{y_A \in A}\left[\sum_{y_B \in B} W_\Sigma(y_B, y_A) \frac{1}{\sigma_i}\left(\frac{(y_B - y_A)_i^2}{\sigma_i^2} - 1\right)\right] , \tag{B.10}$$

where for every $y_A \in A:\ B = A/\{y_A\}$ and $W(y_B, y_A) = \frac{G_\Sigma(y_B - y_A)}{\sum_{y'_B \in B} G_\Sigma(y'_B - y_A)}$. $W(y_B, y_A)$ shows, that evaluating the empirical entropy on sample points from set $A$ the most nearby points from set $B$ give the highest weighted information about the density. In the derivation from equation B.8 to B.9 we used

the simple result that

$$\begin{aligned}\frac{\partial}{\partial\sigma}G_\sigma(x) &= \frac{\partial}{\partial\sigma}\left(\frac{1}{\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right)\right)\\ &= \frac{x^2}{\sigma^3}G_\sigma(x)-\frac{1}{\sigma}G_\sigma(x)=\frac{1}{\sigma}G_\sigma(x)\left(\frac{x^2}{\sigma^2}-1\right)\,. \end{aligned} \tag{B.11}$$

The convergence criteria for the Parzen estimate in our simulations was $\overline{\sum_{i=1}^{D}\Delta\sigma_i} < 0.05$ averaged over the last 40 learning steps.

Note that it is essential for the stochastic learning that the data sample $A$ reflects the real distribution of the random variable. The correct convergence of the learning is assured by re-sampling a large enough $A$ for each learning step. The size of $A$ scales with the dimensionality of the random variable. The probability that one sample falls into a $d$ dimensional sphere with radius of $\sigma$ scales with $\left(\frac{\sigma}{R}\right)^d$ ($R$ is the approximate radius of the area where the data points are distributed). It is "reasonable" to say that $\sigma \ll R$ and therefore the probability decreases exponentially. This means that the estimation for the density becomes highly inaccurate with increasing dimensions unless the number of sample points is not increased exponentially with the number of dimensions. This limits the dimensionality that is still computationally accessible. The computational cost for a learning step is proportional to $d \cdot N_A \cdot N_B$.

### B.2.2 Estimation of the optimal competition parameter

Here we describe the procedure for maximizing the information transfer of the input-output mapping $g(\vec{x})$ for a given output noise level $n(t)$. Note that every learning step for the input-output mapping is based on the Parzen estimate for the output distribution (see the chart of the estimation algorithm). This estimate is determined as it is described in the previous section.

Our original problem in section 4.3 was to estimate the optimal recurrent cortical competition strength. We argued that the "soft-max" function accounts well for the nature of visual cortical processing (see section 4.3.1). The firing rate of a cortical unit $i$ is

$$u_i = g_i(\vec{x},\beta) = \frac{\exp(\beta x_i)}{\sum_j \exp(\beta x_j)}\,, \tag{B.12}$$

where $\beta$ can be interpreted as a competition parameter. Here the mean firing rate in the entire network is assumed to be constant one for a fixation period. Our objective is to find the optimal competition parameter $\beta$ that maximizes the mutual information between input and output for a given noise level. Because the noise here is assumed to be additive, this is equivalent to maximizing entropy of the output.

The entropy of the output $\vec{y} = g(\vec{x},\beta) + n$ is maximized by stochastic gradient ascend on the estimated empirical entropy $h_\Sigma^*[A]$ w. r. t. $\beta$

$$\begin{aligned}\Delta\beta &\propto \frac{\partial}{\partial\beta}h_\Sigma^*[A] = -E_{y_A\in A}\left[\frac{\frac{1}{N_B}\sum_{y_B\in B}\frac{\partial}{\partial\beta}G_\Sigma(y_B-y_A)}{\frac{1}{N_B}\sum_{y_B\in B}G_\Sigma(y_B-y_A)}\right]\\ &= E_{y_A\in A}\left[\frac{\sum_{y_B\in B}G_\Sigma(y_B-y_A)\cdot(y_B-y_A)^{\mathrm{T}}\cdot\Sigma^{-1}\cdot\frac{\partial}{\partial\beta}(y_B-y_A)}{\sum_{y_B\in B}G_\Sigma(y_B-y_A)}\right]\,, \end{aligned} \tag{B.13}$$

where $A$ is a large enough sample from the output. In our simulations the sample size was $|A| = 300$. The gradient on an output component $y_i = g_i(\vec{x}, \beta) + n_i$ (Eq. B.12) is

$$\begin{aligned} \frac{\partial}{\partial \beta} y_i &= \frac{\partial}{\partial \beta} \frac{\exp(\beta x_i)}{\sum_j \exp(\beta x_j)} = \frac{x_i \exp(\beta x_i) \sum_j \exp(\beta x_j) - \exp(\beta x_i) \sum_j x_j \exp(\beta x_j)}{\left(\sum_j \exp(\beta x_j)\right)^2} \\ &= \frac{\exp(\beta x_i) \left(x_i \sum_j \exp(\beta x_j) - \sum_j x_j \exp(\beta x_j)\right)}{\left(\sum_j \exp(\beta x_j)\right)^2} . \end{aligned} \tag{B.14}$$

The gradient of the mapping w. r. t. $\beta$ (Eq. B.14) is independent of the noise because of its additive nature. However, the learning rule (Eq. B.13) depends on the noise reflected in the output $y$. To avoid the calculation of large exponentials in Eq. B.14, for the implementation we take the following equivalent form by dividing both the numerator and denominator by $(\exp(\beta x_i))^2$

$$\frac{\partial}{\partial \beta} \frac{\exp(\beta x_i)}{\sum_j \exp(\beta x_j)} = -\frac{\sum_j (x_j - x_i) \exp\left(\beta (x_j - x_i)\right)}{\left(\sum_j \exp\left(\beta (x_j - x_i)\right)\right)^2} . \tag{B.15}$$

## B.3 Empirical entropy manipulation—Poisson spiking

Now we consider a more realistic Poisson spiking model. We assume that the spike count in increasing intervals beginning with the start of the fixation periods forms the neuronal code. As in the previous section, the visual cortical mapping is modeled by the soft-max function. The expected spike count of a cortical unit $i$ at time $t$ after the beginning of a fixation period is

$$u_i = g_i(\vec{x}, \beta) = \mu(t) \frac{\exp(\beta x_i)}{\sum_j \exp(\beta x_j)} , \tag{B.16}$$

where $\beta$ can be interpreted as the competition parameter and the expected spike count $\mu(t) = r \cdot t$ increases proportionally to the average firing rate $r$ in the entire network.

Here we describe the procedure to determine the competition parameter $\beta$ of the soft-max mapping (Eq. B.12) that maximizes the information transfer for Poisson distributed output noise with a given mean $\mu(t)$. The Poisson noise is multiplicative, therefore we need to maximize the entropy of the output and minimize the conditional uncertainty of the mapping

$$I[X,Y] = H[Y] - H[Y|X] \stackrel{!}{=} \max . \tag{B.17}$$

The realistic mean spike counts within the maximal $200 - 300$ ms long time window are relatively low, with a maximal firing rate of 50 Hz. It follows, that the Poisson distributed output spike count can take only a few values with a considerable probability (we consider less than 15). This allows us to use a direct estimation of the output entropy $H(Y)$ based on the relative frequencies of the different spike counts. Thus, compared to the algorithm for Gaussian distributed output noise, here the outer loop for obtaining a maximum likelihood estimation for the output distribution is omitted (cf. Appendix B.2).

The conditional entropy of the output units for a given input $H(Y|\vec{x})$ is

$$H[Y|\vec{x}] = H[Y|g(\vec{x}, \beta)] = \sum_{i=1}^{D} H[Y|g_i(\vec{x}, \beta)] = \sum_{i=1}^{D} H[Y|u_i] \tag{B.18}$$

because the input-output mapping $g(\vec{x}$ (Eq. B.16) is deterministic and the noise on the output units is statistically independent. The number of dimensions is denoted by $D$. The conditional uncertainty for Poisson noise, given the mean $u$, is (Stein, 1967)

$$H[Y|u_i] = -u_i \log_2\left(\frac{u_i}{e}\right) + e^{-u_i} \sum_{j=2}^{\infty} \frac{log_2(j!)}{j!} u_i^j \ . \tag{B.19}$$

The central limit theorem allows us to approximate the conditional uncertainty of the Poisson process with the uncertainty of the normal distribution with standard deviation $\sqrt{u_i}$ if the expected value is large enough ($u_i > 4$) (Stein, 1967)

$$H[Y|u_i] \approx \log_2\left(\sqrt{2\pi e u_i}\right) \ . \tag{B.20}$$

Using the above approximation for large enough mean values ($u > 4$), we can avoid the calculation of large factorials. The expected uncertainty is

$$H[Y|X] = E_{\vec{x}\in X}\left[\sum_{i=1}^{D} H[Y_i|g_i(\vec{x},\beta)]\right] = D \cdot E_{\vec{x}\in X}\left[H[Y_1|g_1(\vec{x},\beta)]\right] \ , \tag{B.21}$$

because the output distribution is similar for each dimensions $i$.

The mutual information between input $X$ and output $Y$ (Eq. B.17) is maximized by stochastic gradient ascend on the estimated mutual information $I^*[X,Y]$. The learning rule for $\beta$ is

$$\begin{aligned} \Delta\beta \propto \frac{\partial}{\partial\beta} I^*[X,Y] &\approx \frac{H^*_{\beta+\Delta\beta}[Y] - H^*_{\beta}[Y]}{\Delta\beta} - \frac{\partial}{\partial\beta} E_{\vec{x}\in W}\left[\sum_{i=1}^{D} H[Y_i|g_i(\vec{x},\beta)]\right] \\ &= \frac{H^*_{\beta+\Delta\beta}[Y] - H^*_{\beta}[Y]}{\Delta\beta} - D \cdot E_{\vec{x}\in W}\left[\frac{\partial}{\partial\beta} H[Y_1|g_1(\vec{x},\beta)\right] \ , \end{aligned} \tag{B.22}$$

where $W \subset X$ is a large enough sample of the input values $X$ and $H^*_{\beta}[Y]$ is the output entropy at the mapping parameter $\beta$ estimated on the input sample $W$ using the relative frequencies of the different spike counts. The step size $\Delta\beta$ for approximating the derivative of the output entropy must not be too small, because on the fine scale the empirical entropy fluctuates due to the finiteness of the sample.

The derivative of the conditional uncertainty of the Poisson distributed spiking (Eq. B.18) w. r. t. $\beta$ is

$$\begin{aligned} \frac{\partial}{\partial\beta} H[Y|g_i(\vec{x},\beta)] &= -g_i'(\vec{x},\beta) \log_2\left(\frac{g_i(\vec{x},\beta)}{e}\right) - \frac{g_i'(\vec{x},\beta)}{\log(2)} \\ &+ \sum_{j=2}^{\infty} \frac{\log_2(j!)}{j!} g_i'(\vec{x},\beta) \exp(-g_i(\vec{x},\beta))\, g_i(\vec{x},\beta)^{j-1} (j - g_i(\vec{x},\beta)) \\ &= g_i'(\vec{x},\beta)\Big\{ -\log_2\left(\frac{g_i(\vec{x},\beta)}{e}\right) - \frac{1}{log(2)} \\ &+ \exp(-g_i(\vec{x},\beta)) \sum_{j=2}^{\infty} \frac{\log_2(j!)}{j!} g_i(\vec{x},\beta)^{j-1} (j - g_i(\vec{x},\beta)) \Big\} \ , \end{aligned} \tag{B.23}$$

where $g_i'(\vec{x})$ is the derivative of the soft-max mapping w. r. t. $\beta$ (Eq. B.14). For large expected spike counts ($u_i > 4$) the derivative of the approximation (Eq. B.20) of the conditional uncertainty w. r. t. $\beta$ is used.

# Appendix C

# Derivations for the contrast adaptation model in chapter 5

## C.1 Effective transfer function

Here we derive an analytically tractable approximation of the effective transfer function of the model neurons. This transfer function is used only for the derivation of the adaptation rule in section 5.3.1 but not in the simulations. In the following, we omit the indexes $i$ and $j$.

First, instead of determining $R(t)$ for a particular spike train, we assume that the presynaptic spikes are generated by a Poisson process and compute (see appendix C.2) the *expected* amount of synaptic transmitter $\bar{R}(t)$ given an average presynaptic firing rate $f(t)$ (Tsodyks et al., 1998),

$$\frac{d\bar{R}(t)}{dt} = \frac{1-\bar{R}(t)}{\tau_{\text{rec}}} - f(t)p(t)\bar{R}(t) = -\frac{\bar{R}(t)}{\tau_{\text{eff}}(f(t),p(t))} + \frac{1}{\tau_{\text{rec}}}. \tag{C.1}$$

This differential equation describes the evolution of the expected synaptic transmitter with the recovery time constant of $\tau_{\text{rec}} = 200\,\text{ms}$. The effective time constant

$$\tau_{\text{eff}}(f,p) = \tau_{\text{rec}}/(\tau_{\text{rec}} f p + 1) \tag{C.2}$$

for the transmitter dynamics scales inversely with the presynaptic rate and the transmitter release probability. In the derivation of the effective transfer function $O(f,p)$ we assume that the geniculate firing rate $f$ and the release probability $p$ are constant, because the modulation of the geniculate activity (2 Hz drifting grating stimuli that are generally used in experiments exploring contrast adaptation) and the adaptation of the release probability ($\tau_{\text{adapt}} = 7\,\text{sec}$) are slow compared to the transmitter dynamics. The amount of synaptic transmitter changes on a fast time scale of $\tau_{\text{eff}}(f,p) \approx 10\,\text{to}\,200\,\text{ms}$. In this case we can neglect transient terms and approximate the amount of synaptic transmitter for a given presynaptic firing rate and release probability with the steady state solution of equation C.1

$$R^{\infty}(f,p) = \frac{1}{1+fp\tau_{\text{rec}}}\,. \tag{C.3}$$

The stationary EPSC $I^\infty(f,p) \propto f\,p\,R^\infty(f,p)$ is proportional to the presynaptic firing frequency $f$ and the amount of released transmitter. It describes the mean excitatory input to a cortical neuron after an initial transient phase of a few input spikes. This mean current saturates for high input rates $f$ and scales with the transmitter release probability $p$.

To obtain an expression for the output firing rate $O(f,p)$, we have simulated an integrate-and-fire neuron according to equations 5.1, 5.2 and 5.3 driven by Poisson input spike trains. The output rate can be fitted by an effective transfer function

$$O(f,p) \quad \propto \quad S(f)I^\infty(f,p) = \frac{S(f)\,f\,p}{1+f\,p\,\tau_{\text{rec}}}\,, \tag{C.4}$$

where the function $S(f) = \frac{f^\alpha}{f+\Theta}$ accounts for the frequency dependent summation of EPSC pulses with parameter values $\alpha = 0.8$ and $\Theta = 15\,\text{Hz}$. Equation C.4 is the approximation for the input-output mapping realized by a neuron in the homogenous recurrent excitatory network that is used in section 5.3.1 for the derivation of the slow adaptation rule.

## C.2 Mean-field derivation for the synaptic transmitter

Here we derive the dynamics for the expected value of the transmitter $R$. Assuming Poisson presynaptic firing statistics, the expectation $E[.]$ over different realizations of the presynaptic spike train is calculated. From equation 5.3 we obtain

$$E\left[R(T+\Delta)-R(T)\right] \quad = \quad E\left[\int_T^{T+\Delta} \frac{1-R(t)}{\tau_{\text{rec}}}\,dt\right] - E\left[\int_T^{T+\Delta} p(t)R(t)\,dN(t)\right]\,. \tag{C.5}$$

For the derivation three key observations are used. First, the arrival of a spike in the time interval $[t,t+\Delta]$ is independent of $p(t)R(t)$. Second, the dynamics of $p(t)$ is slow compared to $R(t)$ such that $E[p(t)] \approx p(t)$, thus $p(t)$ is approximately independent of $R(t)$. It follows that $E[p(t)R(t)\,dN(t)]$ factorizes. Third, the expectation of the change in the Poisson spike counter is proportional to the presynaptic firing rate: $E[dN(t)] = f(t)dt$. Equation C.1 is then obtained by dividing both sides of equation C.5 with $\Delta$ and taking the limit $\Delta \to 0$.

# Bibliography


Abbott, L. F., Varela, J. A., Sen, K. and Nelson, S. B. (1997), 'Synaptic depression and cortical gain control', *Science* **275**, 220–224.

Adorján, P., Barna, G., Érdi, P. and Obermayer, K. (1998), 'A statistical neural field approach to orientation selectivity', *Neurocomputing* **26-27**, 313–318.

Adorján, P., Levitt, J., Lund, J. and Obermayer, K. (1999), 'A model for the intracortical origin of orientation preference and tuning in macaque striate cortex', *Vis. Neurosci.* **16**, 303–318.

Adorján, P. and Obermayer, K. (1999), Contrast adaptation in simple cells by changing the transmitter release probability, *in* M. S. Kearns, S. A. Solla and D. A. Cohn, eds, 'Advances in Neural Information Processing Systems NIPS 11', MIT Press, Cambridge, Massachusetts, pp. 76–82. In press.

Adorján, P., Piepenbrock, C. and Obermayer, K. (1999), 'Contrast adaptation and infomax in visual cortical neurons', *Rev. Neurosci.* **10**, 181–200. ftp://ftp.cs.tu-berlin.de/pub/local/ni/papers/adp99-contrast.ps.gz.

Adorján, P., Schwabe, L., Piepenbrock, C. and Obermayer, K. (2000), Recurrent cortical competition: Strengthen or weaken?, *in* S. A. Solla, T. K. Leen and K.-R. Mller, eds, 'Advances in Neural Information Processing Systems 12', MIT Press, Cambridge, Massachusetts. In press.

Ahmed, B., Allison, J. D., Douglas, R. J. and Martin, K. A. C. (1997), 'Intracellular study of the contrast-dependence of neuronal activity in cat visual cortex.', *Cerebral Cortex* **7**, 559–570.

Albrecht, D. G. and Hamilton, D. B. (1982), 'Striate cortex of monkey and cat: Contrast response function', *J. Neurophysiol.* **48**, 217–237.

Anderson, J. C., Douglas, R. J., Martin, K. A. and Nelson, J. C. (1994), 'Synaptic output of physiologically identified spiny stellate neurons in cat visual cortex', *J. Comp. Neurol.* **341**, 16–24.

Anderson, J. C., Martin, A. C. and Whitteridge, D. (1993), 'Form, function, and intracortical projections of neurons in the striate cortex of monkey macacus nemestrinus', *Cerebral Cortex* **3**, 412–420.

Antonini, A. and Stryker, M. P. (1993), 'Rapid remodeling of axonal arbors in the visual cortex', *Science* **260**, 1819–1821.

Artun, Ö. B., Shouval, H. Z. and Cooper, L. N. (1998), 'The effect of dynamic synapses on spatiotemporal receptive fields in visual cortex', *Proc. Natl. Acad. Sci.* **95**, 11999–12003.

Atick, J. J. (1992), 'Could information theory provide an ecological theory of sensory processing', *Network* **3**, 213–251.

Atick, J. J. and Redlich, A. N. (1990), 'Towards a theory of early visual processing', *Neural Comput.* **2**, 308–320.

Atick, J. J. and Redlich, A. N. (1992), 'What does the retina know about natural scenes', *Neural Comput.* **4**, 196–210.

Baddeley, R. (1996), 'An efficient code in V1?', *Nature* **381**, 560–561.

Barlow, H. B. (1961), Possible principles underlying the transformation of sensory messages., *in* W. A. Rosenblith, ed., 'Sensory communication', M.I.T. Press, Cambridge, MA, pp. 217–234.

Barlow, H. B. and Földiák, P. (1989), Adaptation and decorrelation in the cortex, *in* R. Durbin, C. Miall and C. Mitchison, eds, 'The computing neuron', Workingham: Addison-Wesley, pp. 54–72.

Barna, G., Grőbler, T. and Érdi, P. (1998), 'Statistical model of the hippocampal CA3 region II. the population framework: model of rhythmic activity in the CA3 slice', *Biol. Cyb.* **79**, 309–321.

Bartsch, H., Stetter, M. and Obermayer, K. (1997), A model for orientation tuning and contextual effects of orientation selective receptive fields, *in* W. Gerstner, A. Germond, M. Hasler and J. Nicoud, eds, 'Artificial Neural Networks – ICANN97', Springer, pp. 237–242.

Batschelet, E. (1981), *Circular statistics in biology*, Academic Press.

Bauer, U., Adorján, P., Scholz, M., Levitt, J., Lund, J. and Obermayer, K. (1997), On the anatomical basis of field size, contrast sensitivity, and orientation selectivity in macaque striate cortex: A model study, *in* W. Gerstner, A. Germond, M. Hasler and J. Nicoud, eds, 'Artificial Neural Networks – ICANN97', Springer, pp. 213–218.

Beaulieu, C., Kisvárday, Z., Somogyi, P., Cynader, M. and Cowey, A. (1992), 'Quantitative distribution of GABA-immunopositive and -immunonegative neurons and synapses in the monkey striate cortex (area 17)', *Cerebral Cortex* **2**, 295–309.

Bell, A. J. and Sejnowski, T. J. (1995), 'An information-maximization approach to blind separation and blind deconvolution', *Neural Comput.* **7**(6), 1129–1159.

Bell, A. J. and Sejnowski, T. J. (1997), 'The 'independent components' of natural scenes are edge filters', *Vision Res.* **37**, 3327–3338.

Ben-Yishai, R., Bar-Or, R. L. and Sompolinsky, H. (1995), 'Theory of orientation tuning in visual cortex', *Proc. Natl. Acad. Sci.* **92**, 3844–3848.

Berman, N. J., Douglas, R. J. and Martin, K. A. C. (1992), GABA-mediated inhibition in the neural networks of visual cortex, *in* R. R. Mize, R. E. Marc and A. M. Sillito, eds, 'Progress in brain research', Vol. 90, Elsevier, New York, pp. 443–476.

Blasdel, G. G. and Fitzpatrick, D. (1984), 'Physiological organization of layer 4 in macaque striate cortex', *J. Neurosci.* **4**, 880–895.

Blasdel, G. G. and Lund, J. S. (1983), 'Termination of afferent axons in macaque striate cortex', *J. Neurosci.* **3**, 1389–1413.

Bloomfield, S. A. (1994), 'Orientation-sensitive amacrine and ganglion cells in the rabbit retina.', *J. Neurophysiol.* **71**, 1672–1691.

Bolz, J. and Gilbert, C. D. (1989), 'The role of horizontal connections in generating long receptive fields in cat visual cortex', *Eur. J. Neurosci.* **1**, 263–268.

Bonds, A. B. (1989), 'Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex', *Vis. Neurosci.* **2**, 41–55.

Bonds, A. B. (1991), 'Temporal dynamics of contrast gain in single cells of cat striate cortex', *Vis. Neurosci.* **6**, 239–255.

Bonhoeffer, T. and Goedecke, I. (1996), 'Development of identical orientation maps for two eyes without common visual experience', *Nature* **379**, 251–4.

Borg-Graham, L. J., Monier, C. and Frégnac, Y. (1998), 'Visual input evokes transient and strong shunting inhibition in visual cortical neurons', *Nature* **393**, 369–373.

Bosking, W. H., Zhang, Y., Schofield, B. and Fitzpatrick, D. (1997), 'Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex', *J. Neurosci.* **16**, 2112–2127.

Brunel, N. and Nadal, J. P. (1997), Optimal tuning curves for neurons spiking according to a poisson process in response to a scalar stimulus, *in* 'Proceedings of ESANN'97'.

Carandini, M., Barlow, H. B., O'Keefe, L. P., Poirson, A. B. and Movshon, J. A. (1997), 'Adaptation to contingencies in macaque primary visual cortex', *Phil. Trans. Roy. Soc. Lond. B* **352**, 1149–1154.

Carandini, M. and Ferster, D. (1997), 'A tonic hyperpolarization underlying contrast adaptation in cat visual cortex', *Science* **276**, 949–952.

Carandini, M., Heeger, D. J. and Movshon, J. A. (1997), 'Linearity and normalization in simple cells of the macaque primary visual cortex', *J. Neurosci.* **17**, 8621–8644.

Carandini, M., Movshon, J. A. and Ferster, D. (1998), 'Pattern adaptation and cross-orientation interactions in the primary visual cortex', *Neuropharmacology* **37**, 501–511.

Carandini, M. and Ringach, D. L. (1997), 'Predictions of a recurrent model of orientation selectivity', *Vision Res.* **37**, 3061–3071.

Celebrini, S., Thorpe, S., Trotter, Y. and Imbert, M. (1993), 'Dynamics of orientation coding in are V1 of the awake monkey', *Vis. Neurosci.* **10**, 811–825.

Chance, F. S., Nelson, S. B. and Abbott, L. F. (1998), 'Synaptic depression and the temporal response characteristics of V1 cells', *J. Neurosci.* **18**, 4785–4799.

Chapman, B., Zahs, K. R. and Stryker, M. P. (1991), 'Relation of cortical cell orientation selectivity to aligment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex', *J. Neurosci.* **11**, 1347–1358.

Croner, L. J. and Kaplan, E. (1995), 'Receptive fields of P and M ganglion cells across the primate retina', *Vision Res.* **35**, 7–24.

Crook, J. M., Kisvárday, Z. F. and Eysel, U. T. (1997), 'GABA-induced inactivation of functionally characterized sites in cat striate cortex: Effects on orientation tuning and direction selectivity', *Vis. Neurosci.* **14**, 141–158.

Crook, J. M., Kisvárday, Z. F. and Eysel, U. T. (1998), 'Evidence for a contribution of lateral inhibition to orientation tuning and direction selectivity in cat visual cortex: Reversible inactivation of functionally characterized sites combined with neuroanatomical tracing techniques', *Eur. J. Neurosci.* **10**, 2056–2075.

Crook, S. M., Ermentrout, G. B. and Bower, J. M. (1998), 'Spike frequency adaptation affects the synchronization properties of networks of cortical oscillations', *Neural Comput.* **10**, 837–854.

Das, A. (1996), 'Orientation in visual cortex: A simple mechanism emerges', *Neuron* **16**, 447–480.

Das, A. and Gilbert, C. D. (1997), 'Distorsions of visuotopic map match orientation singularities in primary visual cortex', *Nature* **387**, 594–598.

De Valois, R. L., Albrecht, D. G. and Thorell, L. G. (1982), 'Spatial frequency selectivity of cells in macaque visual cortex', *Vision Res.* **22**, 545–559.

DeAngelis, G. C., Robson, J. G., Ohzawa, I. and Freeman, R. D. (1992), 'Organization of suppression in receptive fields of neurons in cat visual cortex', *J. Neurophysiol.* **68**, 144–163.

DeBruyn, E. J. and Bonds, A. B. (1986), 'Contrast adaptation in cat visual cortex is not mediated by GABA', *Brain Res.* **383**, 339–342.

Dimitrov, A. and Cowan, J. D. (1998), 'Spatial decorrelation in orientation-selective cortical cells', *Neural Comput* **10**, 1779–1795.

Dinse, H. R., Kruger, K., Mallot, H. A. and Best, J. (1991), Temporal structure of cortical information processing: Cortical architecture, oscillations, and non-separability of spatio-temporal receptive field organization, *in* J. Kruger, ed., 'Neural Cooperativity', Springer, pp. 68–104.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C. and Suarez, H. H. (1995), 'Recurrent excitation in neocortical circuits', *Science* **269**, 981–985.

Douglas, R. J., Martin, K. A. C. and Whitteridge, D. (1988), 'Selective responses of visual cortical neurons do not depend on shunting inhibition', *Nature* **332**, 642–644.

Douglas, R. J., Martin, K. A. C. and Whitteridge, D. (1991), 'An intracellular analysis of the visual responses of neurons in cat visual cortex', *J. Physiol.* **440**, 659–696.

Douglas, R. J. and Martin, K. C. (1991), 'A functional microcircuit for cat visual cortex', *J. Physiol.* **440**, 735–769.

Engel, A. K., Roelfsema, P. R., Fries, P., Brecht, M. and Singer, W. (1997), 'Role of the temporal domain for response selection and perceptual binding', *Cerebral Cortex* **7**, 571–582.

Eysel, U. T., Crook, J. M. and Machemer, H. F. (1990), 'GABA-induced remote inactivation reveals cross-orientation inhibition in the cat striate cortex', *Exp. Brain Res.* **80**, 626–630.

Ferster, D. (1986), 'Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex', *J. Neurosci.* **6**, 1284–1301.

Ferster, D., Chung, S. and Wheat, H. (1996), 'Orientation selectivity of thalamic input to simple cells of cat visual cortex', *Nature* **380**, 249–281.

Ferster, D. and Jagadeesh, B. (1992), 'EPSP-IPSP interactions in cat visaual cortex studied in vivo whole-cell patch recording', *J. Neurosci.* **12**, 1262–1274.

Field, D. J. (1987), 'Relations between the statistics of natural images and the response properties of cortical cells', *J. Opt. Soc. Am. A* **4**, 2379–2394.

Field, D. J. (1994), 'What is the goal of sensory coding', *Neural Comput.* **6**, 559–601.

Finlayson, P. G. and Cynader, M. S. (1995), 'Synaptic depression in visual cortex tissue slices: an in vitro model for cortical neuron adaptation', *Exp. Brain Res.* **106**, 145–155.

Fitzpatrick, D., Lund, J. S., Schmechel, D. E. and Towles, A. C. (1987), 'Distribution of GABAergic neurons and axon terminals in the macaque striate cortex', *J. Comp. Neurol.* **264**, 73–91.

Freund, T. F., Martin, K. A. C., Soltesz, I., Somogyi, P. and Whitteridge, D. (1989), 'Arborisation pattern and postsynaptic targets of physiologically identified thalamocortical afferents in striate cortex of macaque monkey', *J. Comp. Neurol.* **289**, 315–336.

Gibson, J. J. (1966), *The perception of the visual world*, Houghton Mifflin, Boston.

Grőbler, T., Barna, G. and Érdi, P. (1998), 'Statistical model of the hippocampal CA3 region I. single cell and population activity in a statistical model of the hipocampal CA3 region', *Biol. Cyb.* **79**, 301–308.

Hammond, P. and Pomfrett, C. J. D. (1989), 'Influence of spatial frequency on tuning and bias for orientation and direction in the cat's striate cortex', *Vision Res.* **30**, 359–369.

Hata, Y., Tsumoto, T., Sato, H., Hagihara, K. and Tamura, H. (1988), 'Inhibition contributes to orientation selectivity in visual cortex of cat.', *Nature* **335**, 815–817.

Hawken, M. J. and Parker, A. J. (1984), 'Contrast sensitivity and orientation selectivity in lamina IV of the striate cortex of Old World monkeys', *Exp. Brain Res.* **54**, 367–372.

Hawken, M. J. and Parker, A. J. (1987), 'Spatial properties in the monkey striate cortex', *Proc. Roy. Soc. Lond. B* **231**, 251–288.

Heller, J., Hertz, J. A., Kjaer, T. W. and Richmond, B. J. (1995), 'Information flow and temporal coding in primate pattern vision', *J. Comput. Neurosci.* **2**, 175–193.

Holt, G., Softky, W. R., Koch, C. and Douglas, R. J. (1996), 'A comparison of discharge variability in vitro and in vivo in cat visual cortical neurons', *J:Neurophys* **75**, 1806–1814.

Hubel, D. H. and Wiesel, T. N. (1962), 'Receptive fields, binocular interaction and functional architecture in cat's visual cortex', *J. Physiol.* **165**, 559–568.

Hubel, D. H. and Wiesel, T. N. (1974), 'Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor', *J. Comp. Neur.* **158**, 295–306.

Jagadeesh, B. and Ferster, D. (1990), 'Receptive field lengths in cat striate cortex can increase with decreasing stimulus contrast', *Society of Neuroscience Abstracts* **16**, 293.

Jones, J. P. and Palmer, L. A. (1987*a*), 'An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex', *J. Neurophysiol.* **58**, 1233–1258.

Jones, J. P. and Palmer, L. A. (1987*b*), 'The two-dimensional spatial structure of simple receptive fields in cat striate cortex', *J. Neurophysiol.* **58**, 1187–1211.

Jones, J. P., Stepnoski, A. and Palmer, L. A. (1987), 'The two-dimensional spectral structure of simple receptive fields in cat striate cortex', *J. Neurophysiol.* **58**, 1212–1232.

Kandel, E. R., Schwartz, J. H. and Jessell, T. M. (1991), *Principles of neural science*, Prentice-Hall, London.

Kaplan, E., Purpura, K. and Shapley, R. M. (1987), 'Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus', *J. Physiol.* **391**, 267–288.

Kim, D. S. and Bonhoeffer, T. (1994), 'Reverse occlusion leads to a precise restoration of orientation preference maps in visual cortex', *Nature* **370**, 370–372.

Komatsu, Y., Nakajima, S., Toyama, K. and Fetz, E. (1988), 'Intracortical connectivity revealed by spike-triggered averaginng in slice preparations of cat visual cortex', *Brain Res* **442**, 359–362.

König, P. and Schillen, T. B. (1991), 'Stimulus-dependent assembly formation of oscillatory responses: I. synchronization and II. desynchronization', *Neural Comput.* **3**, 155–177.

Laughlin, S. B. (1981), 'A simple coding procedure enhances a neuron's information capacity', *Z. Naturforschung* **36**, 910–912.

Laughlin, S. B. (1994), 'Matching coding, circuits, cells, and molecules to signals: general principles of retinal design in the fly's eye', *Prog. Ret. Eye Res.* **13**, 165–196.

Lee, D. K., Itti, L., Kock, C. and Braun, J. (1999), 'Attention activates winner-take-all competition among visual filters', *Nat. Neurosci.* **2**, 375–381.

Leopold, D. A. and Logothetis, N. K. (1998), 'Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex', *Exp. Brain Res.* **123**, 341–345.

Levitt, J. B. and Lund, J. S. (1997), 'Contrast dependence of contextual effects in primate visual cortex', *Nature* **387**, 73–76.

Linsker, R. (1989), 'How to generate ordered maps by maximizing the mutual information between input and output signals.', *Neural Comput.* **1**, 402–411.

Livingstone, M. S. (1998), 'Mechanisms of direction selectivity in macaque v1', *Neuron* **20**, 509–526.

Lund, J. S. (1980), Intrinsic organization of the primate visual cortex, area 17, as seen in Golgi preparations, *in* E. Adelman, F. I. Schmidt, F. G. Worden and S. G. Dennis, eds, 'The organization of cerebral cortex: Proceedings of neuroscience research program colloquium', MIT Press Cambridge, Massachusetts, pp. 105–124.

Lund, J. S. (1987), 'Local circuit neurons of macaque monkey striate cortex: I. Neurons of laminae 4C and 5A', *J. Comp. Neurol.* **257**, 60–92.

McCormick, D. A., Connors, B. W., Lighthall, J. W. and Prince, D. A. (1985), 'Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex', *J. Neurophysiol.* **54**, 782–806.

McLean, J. and Palmer, L. A. (1996), 'Contrast adaptation and excitatory amino acid receptors in cat striate cortex', *Vis. Neurosci.* **13**, 1069–1087.

Mechler, F., Victor, J. D., Purpura, K. P. and Shapley, R. (1998), 'Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli', *J. Neurosci.* **18**, 6583–6598.

Movshon, J. A. and Lennie, P. (1979), 'Pattern selective adaptation in visual cortical neurons', *Nature* **278**, 850–852.

Mundel, T., Dimitrov, A. and Cowan, J. D. (1997), Visual cortex circuitry and orientation tuning, *in* M. Mozer, M. Jordan and T. Petsche, eds, 'Advances in Neural Information Processing Systems', Vol. 9, MIT Press Cambridge, Massachusetts, pp. 887–893.

Nadal, J. P., Brunel, N. and Parga, N. (1998), 'Nonlinear feedforward networks with stochastic ouputs: infomax implies redundancy reduction', *Network: Computation in Neural Systems* **9**, 207–217.

Nelson, S., Toth, L., Sheth, B. and Sur, M. (1994), 'Orientation selectivity of cortical neurons during intracellular blockade of inhibition', *Science* **265**, 774–777.

Nykamp, D. and Tranchina, D. (2000), 'A population-density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning', *J. Comput. Neurosci.* **8**, 19–50.

Ohzawa, I., Sclar, G. and Freeman, R. D. (1985), 'Contrast gain control in the cat's visual system', *J. Neurophysiol.* **54**, 651–667.

Olshausen, B. A. and Field, D. J. (1996), 'Emergence of simple-cell receptive field properties by learning a sparse code for natural images', *Nature* **381**, 607–609.

Olshausen, B. A. and Millman, K. J. (2000), Learning sparse codes with a mixture-of-gaussians prior, *in* S. A. Solla, T. K. Leen and K.-R. Mller, eds, 'NIPS12', MIT Press, Cambridge, Massachusetts. In press.

Pei, X., Vidyasagar, T. R., Volgushev, M. and Creutzfeld, O. D. (1994), 'Receptive field analysis and orientation selectivity of postsynaptic potentials of simple cells in cat visual cortex', *J. Neurosci.* **14**, 7130–7140.

Pernberg, J., Volgushev, M. and Eysel, U. T. (1998), Comparsion of selectivity of postsynaptic potentials and of spike responses in cat visual cortex, *in* 'Eur. J. Neurosci.', Vol. 10, p. 243.

Peters, A., Payne, B. R. and Budd, J. (1994), 'A numerical analysis of the geniculocortical input to striate cortex in the monkey', *Cerebral Cortex* **4**, 215–29.

Piepenbrock, C. and Obermayer, K. (1999), The role of lateral cortical competition in ocular dominance development, *in* M. S. Kearns, S. A. Solla and D. A. Cohn, eds, 'Advances in Neural Information Processing Systems NIPS 11', MIT Press, Cambridge, Massachusetts, pp. 139–145.

Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T. and Norcia, A. M. (1998), 'Collinear stimuli regulate visual responses depending on cell's contrast threshold.', *Nature* **391**, 580–584.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1994), *Numerical recipies in C*, Cambridge University Press.

Reid, R. C. and Alonso, J. M. (1995), 'Specificity of monosynaptic connections from thalamus to visual cortex', *Nature* **378**, 281–284.

Reid, R. C., Soodak, R. E. and Shapley, R. M. (1991), 'Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex', *J. Neurophysiol.* **66**, 505–529.

Richmond, B. J. and Optican, L. M. (1990), 'Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex II. Information transmission', *J. Neurophysiol.* **64**(2), 370–380.

Rieke, F., Warland, D., de Ryter van Steveninck, R. and Bialek, W. (1997), *Spikes: exploring the neural code*, Bradford Books: MIT Press, Cambridge, MA.

Ringach, D. L., Hawken, M. J. and Shapley, R. (1997*a*), 'Dynamics of excitatory and inhibitory mechanisms shaping the orientation tuning of neurons in V1', *Society of Neuroscience Abstracts* **23**, 1544.

Ringach, D. L., Hawken, M. J. and Shapley, R. (1997*b*), 'Dynamics of orientation tuning in macaque primary visual cortex', *Nature* **387**, 281–284.

Rodieck, R. W. (1965), 'Quantitative analysis of cat retinal ganglion cell response to visual stimuli', *Vision Res.* **5**, 583–601.

Ruderman, D. L. (1994), 'The statistics of natural images', *Network: computation in neural systems* **5**, 517–548.

Sabatini, S. P. (1996), 'Recurrent inhibition and clustered connectivity as a basis for Gabor-like receptive fields in the visual cortex', *Biol. Cyb.* **74**, 189–202.

Sato, H., Katsuyama, N., Tamura, H., Hata, Y. and Tsumoto, T. (1996), 'Mechanisms underlying orientation selectivity of neurons in the primary visual cortex of macaque', *J. Physiol.* **494**, 757–771.

Saul, A. B. (1995), 'Adaptation aftereffects in single neurons of cat visual cortex: response timing is retarded by adapting', *Vis. Neurosci.* **12**, 191–205.

Saul, A. B. and Cynader, M. S. (1989*a*), 'Adaptation in single units in visual cortex: the tuning of aftereffects in the spatial domain', *Vis. Neurosci.* **2**, 593–607.

Saul, A. B. and Cynader, M. S. (1989*b*), 'Adaptation in single units in visual cortex: the tuning of aftereffects in the temporal domain', *Vis. Neurosci.* **2**, 609–620.

Schiller, P. (1982), 'Central connections of the retinal ON and OFF pathways', *Nature* **297**, 580–583.

Schumer, R. A. and Movshon, J. A. (1984), 'Length summation in simple cells of cat striate cortex.', *Vision Res.* **24**, 565–571.

Schwabe, L. (1999), Dynamic properties of recurrent biological neural networks, Master's thesis, Technical University Berlin, Dept. of Comp. Sci.

Schwabe, L., Adorján, P. and Obermayer, K. (2000), Spike-frequency adaptation as a mechanism for dynamic coding in V1. Sbumitted.

Schwartz, C. and Bolz, J. (1991), 'Functional specificity of a long-range horizontal connection in cat visual cortex: A cross-correlation study', *J. Neurosci.* **11**, 2995–3007.

Sclar, G. and Freeman, R. D. (1982), 'Invariance of orientation tuning with stimulus contrast', *Exp. Brain Res.* **46**, 457–461.

Sclar, G., Lennie, P. and DePriest, D. D. (1989), 'Contrast adaptation in the striate cortex of macaque', *Vision Res.* **29**, 747–755.

Senn, W., Segev, I. and Tsodyks, M. (1998), 'Reading neural synchrony with depressing synapses', *Neural Comput.* **10**, 815–819.

Senn, W., Wyler, K., Streit, J., Larkum, M., Lüscher, H.-R., H. Mey, L. M. a. D. S., Vogt, K. and Wannier, T. (1996), 'Dynamics of a random neural network with synaptic depression', *Neural Networks* **9**, 575–588.

Shannon, C. E. and Weaver, W. (1949), *The mathematical theory of communication*, University of Illinois Press, Urbana.

Shapley, R. M. and Enroth-Cugell, C. (1984), 'Visual adaptation and retinal gain controls', *Prog. Ret. Eye Res.* **3**, 263–346.

Sherk, H. and Horton, J. C. (1984), 'Receptive field properties in cat's area 17 in absence of ON-center geniculate input', *J. Neurosci.* **4**, 381–393.

Shou, T., Li, X., Zhou, Y. and Hu, B. (1996), 'Adaptation of visually evoked responses of relay cells in the dorsal lateral geniculate nucleus of the cat following prolonged exposure to drifting gratings', *Vision Res.* **13**, 605–613.

Sillito, A. M. (1975), 'The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat', *J. Physiol.* **250**, 305–329.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. and Davis, J. (1995), 'Visual cortical mechanisms detecting focal discontinuities.', *Nature* **378**, 492–496.

Sillito, A. M., Kemp, J. A., Milson, J. A. and Berardi, N. (1980), 'A re-evaluation of the mechanisms underlying simple cell orientation selectivity', *Brain Res.* **194**, 517–520.

Smith, E. L., Chino, Y. M., Ridder, W. H., Kitagawa, K. and Langston, A. (1990), 'Orientation bias of neurons in the lateral geniculate nucleus of macaque monkeys', *Vis. Neurosci.* **5**, 525–45.

Somers, D. C., Nelson, S. B. and Sur, M. (1995), 'An emergent model of orientation selectivity in cat visual cortical simple cells.', *J. Neurosci.* **15**, 5448–65.

Sompolinsky, H. and Shapley, R. (1997), 'New perspectives on the mechanisms for orientation selectivity', *Curr. Op. in Neurobiol.* **7**, 514–522.

Spear, P. D., Moore, R. J., Kim, C. B. Y., Xue, J. T. and Tumosa, N. (1994), 'Effects of aging on the primate visual system: spatial and temporal processing by lateral geniculate neurons in young adult and old rhesus monkeys', *J. Neurophysiol.* **72**, 402–420.

Stein, R. B. (1967), 'The information capacity of nerve cells using a frequency code', *Biophysical Journal* **7**, 797–826.

Stemmler, M. and Koch, C. (1999), Information maximization in single neurons, *in* M. S. Kearns, S. A. Solla and D. A. Cohn, eds, 'Advances in Neural Information Processing Systems NIPS 11', MIT Press. in press.

Stetter, M., Adorján, P., Bartsch, H. and Obermayer, K. (1998), Modelling contrast adaptation and contextual effects in primary visual cortex, *in* 'The Fifth International Conference on Neural Information Processing, ICONIP'98-Kitakyushu', Vol. 2, pp. 669–672.

Stetter, M., Sendtner, R. A. and Timberlake, G. T. (1996), 'A novel method for measuring saccade profiles using the scanning laser ophthalmoscope', *Vision Res.* **36**, 1987–1994.

Stratford, K. J., Tarczy-Hornoch, K., Martin, K. A. C., Bannister, N. J. and Jack, J. (1996), 'Excitatory synaptic inputs to spiny stellate cells in cat visual cortex', *Nature* **382**, 258–261.

Sugase, Y., Yamane, S., Ueno, S. and Kawano, K. (1999), 'Global and fine information coded by single neurons in the temporal visual cortex', *Nature* **400**, 869–873.

Tanaka, K. (1983), 'Cross-correlation analysis of geniculostriate neuronal relationships in cats', *J. Neurophysiol.* **49**, 1303–1318.

Todorov, E., Siapas, A. and Somers, D. (1997), A model of recurrent interactions in primary visual cortex, *in* M. C. Mozer, M. I. Jordan and T. Petsche, eds, 'Advances in Neural Information Processing Systems', Vol. 9, MIT Press Cambridge, Massachusetts, pp. 118–124.

Todorov, E. V., Siapas, A. G., Somers, D. C. and Nelson, S. B. (1997), Modeling visual cortical contrast adaptation effects, *in* J. Bower, ed., 'Computational Neuroscience: Trends in Research', Plenum Press, New York, pp. 525–531.

Troyer, T. W., Krukowski, A. E., Priebe, N. J. and Miller, K. D. (1998), 'Contrast-invariant orientation tuning in visual cortex: Feedforward tuning and correlation-based intracortical connectivity', *J. Neurosci.* **18**, 5908–5927.

Tsodyks, M., Pawelzik, K. and Markram, H. (1998), 'Neural networks with dynamic synapses', *Neural Comput.* **10**, 821–835.

Tsodyks, M. V. and Markram, H. (1997), 'The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability', *Proc. Natl. Acad. Sci.* **94**, 719–723.

Tsodyks, M. V. and Sejnowski, T. (1995), 'Rapid state switching in balanced cortical network models', *Network: Computation in Neural Systems* **6**, 111–124.

Tsumoto, T., Eckart, W. and Creutzfeld, O. D. (1979), 'Modifications of orientation sensitivity of cat visual cortex neurons by removal of GABA-mediated inhibition', *Exp. Brain Res.* **34**, 351–363.

van der Schaaf, A. and van Hateren, J. H. (1996), 'Modelling the power spectra of natural images: Statistics and information', *Vision Res.* **36**, 2759–2770.

Ventriglia, F. (1974), 'Kinetic approach to neural systems', *Bull. Math. Biol.* **52**, 397–429.

Vidyasagar, T. R. (1990), 'Pattern adaptation in cat visual cortex is a co-operative phenomenon', *Neurosci.* **36**, 175–179.

Vidyasagar, T. R., Pei, X. and Volgushev, M. (1996), 'Multiple mechanisms underlying the orientation selectivity of visual cortical neurones', *TINS* **19**, 272–277.

Vidyasagar, T. R. and Sigüenza, J. A. (1985), 'Relationship between orientation tuning and spatial frequency in neurones of cat area 17', *Exp. Brain Res.* **57**, 628–631.

Viola, P., Schraudolph, N. N. and Sejnowski, T. J. (1996), Empirical entropy manipulation for real-world problems, *in* D. S. Touretzky, M. C. Mozer and M. E. Hasselmo, eds, 'Advances in Neural Information Processing Systems 8', MIT Press, Cambridge, MA, pp. 851–857.

Viviani, P. (1990), 'Eye movements in visual search: cognitive, perceptual and motor control aspects', *Rev. Oculomot. Res.* **4**, 353–393.

Volgushev, M., Vidyasagar, T. R. and Pei, X. (1995), 'Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex', *Vis. Neurosci.* **12**, 621–628.

Volgushev, M., Vidyasagar, T. R. and Pei, X. (1996), 'A linear model fails to predict orientation selectivity of cells in the cat visual cortex.', *J. Physiol.* **496**, 597–606.

von der Malsburg, C. P. and Buhmann, J. (1992), 'Sensory segmentation with coupled oscillators', *Biol. Cyb.* **67**, 233–242.

Webster, M. A. and De Valois, R. L. (1985), 'Relationship between spatial-frequency and orientation tuning of striate cells', *J. Opt. Soc. Am. A* **2**, 1124–1132.

Wehmeier, U., Dong, D., Koch, C. and Van Essen, D. (1989), Modeling the visual system., *in* C. Koch and I. Segev, eds, 'Methods in neuronal modeling.', MIT Press, Cambridge, pp. 335–359.

Wiener, M. C., Oram, M. W. and Richmond, B. J. (1999), 'Lantency is a better temporal code than principal components in V1', *Society of Neuroscience Abstracts* **25**, 1549.

Wolf, F., Bauer, H. U., Pawelzik, K. and Geisel, T. (1996), 'Organization of the visual cortex', *Nature* **382**, 306–307.

Wörgötter, F. and Eysel, U. T. (1991), 'Topographical aspects of intracortical excitation and inhibition contributing to orientation specificity in area 17 of the cat visual-cortex', *Eur. J. Neurosci.* **3**, 1232–1244.

Wörgötter, F. and Koch, C. (1991), 'A detailed model of the primary visual pathway in the cat: Comparison of afferent excitatory and intracortical inhibitory schemes for orientation selectivity', *J. Neurosci.* **11**, 1959–1979.

Yoshioka, T., Levitt, J. B. and Lund, J. S. (1994), 'Independence and merger of thalamocortical channels within macaque monkey primary visual cortex: Anatomy of interlaminar projections', *Vis. Neurosci.* **11**, 467–489.

Zipser, K., Lamme, V. A. F. and Schiller, P. H. (1996), 'Contextual modulation in primary visual cortex', *J. Neurosci.* **15**, 7376–7389.